THE INVERTIBILITY OF DYNAMIC SYSTEMS WITH APPLICATION TO CONTROL

ROGER W. BROCKET

SYSTEMS RESEARCH CENTER
CASE INSTITUTE OF TECHNOLOGY
CLEVELAND, OHIO

MAY, 1963

SRC 31-A-63-14

THE INVERTIBILITY OF DYNAMIC SYSTEMS WITH APPLICATION TO CONTROL

Roger W. Brockett
NDEA Fellow

Systems Research Center
Case Institute of Technology
Cleveland, Ohio

May, 1963

SRC 31-A-63-14

ABSTRACT

Certain properties of multi-input, multi-output dynamic systems which are of interest in automatic control are investigated. It is shown that such systems can be viewed as a mapping whose domain is the set of possible input functions and whose range is either the set of responses or a set of equivalence classes. By selecting the norms for the input and output spaces in various ways it is possible to interpret many of the familiar properties of a system in terms of these mapp..gs and new means of system characterization are suggested as well. It is shown that the study of these mappings leads naturally to the study of an inverse equation. Conditions under which the inverse equation exists are derived for some linear and nonlinear systems and explicit representations for the inverse equation are given for certain classes of linear systems.

The effect of feedback is analyzed in terms of its influence on the character of these mappings and certain limitations on feedback as a device for altering a given system are noted. It is also shown that the inverse equation can be used to help define the optimum input for a certain type of time-optimal problems where the objective is stated in terms of the outputs rather than the state. Examples are given which show that the time-optimal forcing function cannot always be generated by an ideal relay.

FOREWORD

The Systems Research Center is a research and graduate study center operating in direct cooperation with all departments and divisions of Case Institute of Technology. The center brings together faculty and students in a coordinated program of research and education in the important techniques of systems theory, development, and application.

Research leading to this report was carried on by Mr. Roger W. Brockett, Graduate Assistant, under the direction of Dr. Mihajlo D. Mesarovic, Associate Professor of Engineering at Case and Director of the Adaptive and Self-Organizing Systems group of the Systems Research Center.

The research activities of this group was supported in part by the Department of Health, Education, and Welfare and the United States Office of Naval Research (Contract NONR 1141(12)).

Ellis A. Johnson, Director
Systems Research Center

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

1.1 The General Problem

A basic property of those systems which are of interest to automatic control engineers is that their outputs can be altered to meet changing requirements by means of input manipulation. It is also true, however, that for most systems there are definite limitations on the types of responses that can be obtained. For example, it is usually impossible to make a physical variable change its value instantaneously. In general, the types of responses that can be obtained from a system depend on the class of possible inputs, the present state of the system, and the nature of the equations governing the system. The purpose of this research is to study the nature of these dependencies, to characterize certain classes of systems which have the property of being able to generate all responses in a given set, and to indicate certain applications of these ideas.

The systems which we will consider are all special cases of those which can be described by a pair of equations of the form

$$\dot{Z}(t) = F(Z(t),X(t)) \tag{1.1a}$$

$$Y(t) = G(Z(t),X(t)) \tag{1.1b}$$

The symbol $X(t)$ denotes the value of the input at time t and $Y(t)$ denotes the corresponding value of the output. The vector $Z(t)$, which is usually called the state, is not assumed to bear any particular relationship to the input or output variables. It should be thought of as being introduced only to enable the first order

representation of the differential equation. The inputs and outputs are permitted to be vector valued and initially are not assumed to be of the same dimension although this will be by far the most important case for our work.

If we are interested in the behavior of the system (1.1) over a specific interval of time, say $[o,\sigma]$, then it is not the value of the variable $X(t)$ at any one time which determines the response but rather it is the behavior of $X(t)$ over the entire interval. To avoid confusion it is imparative to distinguish between a function, that is, a complete description of how the variable varies with time, and the value of a function at some particular time t. As is common in mathematical literature we will use symbols such as X, Y, and Z without arguments to denote functions and use $X(t)$, $Y(t)$, and $Z(t)$ to denote the value of a function at time t.

In terms of this notation the problem under consideration can be given a more precise statement. Let S_x denote a set of functions defined on the interval $[o,\sigma]$. Assume that for each* $X \varepsilon S_x$ the equation (1.1a) has a unique solution passing through a given initial point $Z(0)$ and assume that this solution is defined over the entire interval $[o,\sigma]$. This fact may be expressed by saying that the system (1.1) associates a unique Y with each X in S_x or by saying that (1.1) defines a (single valued) mapping whose domain is S_x. The problem we are studying, that is, the problem of determining the types of responses the system is capable of, is identical to that of

*The symbol ε should be read "belonging to".

determining the range of this mapping.

From an abstract point of view our approach will be to introduce norms in the input and output spaces and then to discuss the conditions under which certain transformations are continuous and invertible. The conditions under which a transformation has a continuous inverse will be particularly important since this corresponds to the case where a more or less arbitrary output can be realized.

Because our approach and methods are somewhat different from those currently in vogue in control systems theory an effort has been made to relate the properties under consideration to the more familiar concepts of stability, minimum phasity, controllability, singularity of transfer matrices, etc. The remainder of this chapter is devoted to a somewhat informal explanation of the approach to be taken. It includes a brief account of our results as well as some of the results of other workers who have made similar investigations or have otherwise contributed to this study.

Although this thesis represents an engineering study it has been necessary to use some mathematical terminology in order to make the intuitive ideas precise. This is particularly true in Chapter II where the objective is to develop the essential points in such a way as to make them acceptable and understandable to both control engineers and applied mathematicians.

1.2 Systems as Transformations

As indicated above, once $Z(0)$ is fixed the system (1.1) can be viewed as a transformation whose domain is a set of functions S_x.

For many purposes it is most natural to regard the range of such a transformation as being the set of response functions, i.e., the set S_y consisting of all the functions Y which can be generated by an X in S_x. If this point of view is adopted then one way of determining what types of responses are possible is to construct a representation of the inverse mapping. For example, if X and Y are related by a system of the form

$$\dot{Z}(t) = AZ(t) + BX(t) \tag{1.2a}$$

$$Y(t) = CZ(t) + DX(t) \tag{1.2b}$$

and if D is nonsingular, then it is possible to "solve" (1.2) for X(t) in terms of Y(t) and Z(t). That is, we can eliminate X(t) from the differential equation by expressing it in terms of Y(t) and Z(t). For the system (1.2) the inverse system is given by

$$\dot{Z}(t) = (A-BD^{-1}C)Z(t) + BD^{-1}Y(t) \tag{1.3a}$$

$$X(t) = D^{-1}Y(t) - D^{-1}CZ(t) \tag{1.3b}$$

Both (1.2) and (1.3) provide complete descriptions of the relationship between X and Y. The important difference is that the representation (1.2) places in evidence the differential equation relating Z to X whereas (1.3) places in evidence the differential equation relating Z to Y. Since X is the independent variable the representation afforded by (1.2) is more natural in most cases and hence we will take it to be the definitive relationship and will regard (1.3) as the inverse transformation.

The importance of the inverse system is twofold. First, the

very fact that it is possible to derive such a relationship implies that for any Y such that (1.3) has a unique solution there exists an X which generated it and in addition (1.3) gives a means for computing it. Second, the representation (1.3) allows one to study how X varies as a function of Y. This is particularly important in cases where the systems of equations cannot be solved and qualitative methods must be used.

In view of the large amount of work which has been done on Liapunov stability it is natural to ask how instability manifests itself if one adopts the proposed point of view. In particular, what is the significance of an unstable inverse equation?

In order to answer these questions satisfactorily it is necessary to examine the concept of continuity as it applies to mappings between function spaces.

1.3 Continuity

Loosely speaking what we have in mind is this. If the eigenvalues of A have negative real parts, and if X is a particular function which maps into Y, then any X' which is "close" to X maps into a function Y' which is "close" to Y. This is true even if we are interested in the behavior of Y(t) over the entire interval $[0,\infty]$. It is not necessarily true, however, that any Y' which is "close" to Y can be generated by an X' which is "close" to the X which generated Y. This is true only if the eigenvalues of the matrix $(A-BD^{-1}C)$ appearing in the inverse equation have negative real parts.

At this point these statements lack precision because it has not been specified what is meant by "close". For the scalars $x(t)$ and $y(t)$, the absolute value of the difference, $|x(t)-y(t)|$, is a measure of closeness. For vectors a suitable definition is the maximum over all i of $|x_i(t)-y_i(t)|$ where $x_i(t)$ and $y_i(t)$ are the $i^{\underline{th}}$ components of $X(t)$ and $Y(t)$ respectively. This will be written as $|X(t)-Y(t)|$. For vector valued functions defined over $[0,\sigma]$ a measure of proximity may be defined as the maximum over all $t\epsilon[0,\sigma]$ of $|X(t)-Y(t)|$. This is denoted by $||X-Y||_\sigma$ or simply by $||X-Y||$ if the interval of interest is $[0,\infty]$. A function such as this which satisfies certain technical requirements is known as a norm. This is by no means the only possible way of defining a norm and although this particular norm will always be used for the input space S_x it will be necessary to consider other types of norms for the output set in order to accurately characterize the transformations under consideration.

Return now to the system (1.2). Let $S_x = C$, the class of continuous functions defined on $[0,\infty]$, and suppose the above definition of norm is used. Then, as will be shown in Chapter II, if the eigenvalues of A have negative real parts then the system (1.2) defines a continuous mapping of C into C. That is to say, if the eigenvalues of A have negative real parts then for any given $\epsilon > 0$ there exists a $\delta > 0$ such that if $||X-X'|| < \delta$ then $||Y-Y'|| < \epsilon$. Here again Y is the image of X and Y' is the image of X'. The condition that the eigenvalues of A have negative real parts is necessary and sufficient for the asymptotic stability of the

differential equation (1.2a) and thus for this particular transformation asymptotic stability implies continuity. It is important and useful to know that the total stability theorem[19*] implies that this is also true for a wide class of time-invariant nonlinear systems, provided that Z(0) is sufficiently close to the critical point and that the inputs are sufficiently small.

The inverse equation (1.3) defines a continuous mapping of C into C if the eigenvalues of $(A-BD^{-1}C)$ have negative real parts. If, on the other hand, there exist eigenvalues of $(A-BD^{-1}C)$ with positive real parts then there may be output functions which are finite but which require an infinite input to produce them. In terms of the approach taken here such a system would be described as one which does not have a continuous inverse.

Another reason that the inverse may fail to be continuous in terms of the characterization of the response space given above is that the inverse system may involve derivatives of Y. Consider the scalar equation

$$\mathring{y}(t) + by(t) = x(t) \tag{1.4}$$

If b is positive then when viewed as a mapping of S_x into S_y (1.4) is continuous provided both S_x and S_y are taken to be C. Difficulties arise, however, when considering the inverse mapping. The fact is that by making $||y-y'||$ small one cannot insure that

*Superscripts refer to the references found immediately following Chapter V.

the inverse images of y and y' will be close. Thus (1.4) does not have a continuous inverse if $S_y = C$.

The problem in this case is that the value of $x(t)$ depends on $\dot{y}(t)$ as well as $y(t)$ whereas a restriction in the value of $||y-y'||$ gives no control over the value of $||\dot{y}-\dot{y}'||$. The solution to this problem is to define a norm which gives control over both $||y-y'||$ and $||\dot{y}-\dot{y}'||$. It should be clear that in more general situations control over higher derivatives may be required as well. Let C^k be the set of functions defined on $[o,\infty]$ which have k continuous derivatives. Define $||X||^k$, the $\mathbb{C}^k$ norm, as maximum over all i from zero to k of $||X^{(i)}||$ where $X^{(i)}$ is the $i^{\underline{th}}$ derivative of X with respect to time. In terms of this notation equation (1.4) defines continuous transformation of $\mathbb{C}$ into $\mathbb{C}^1$ and as such, it has a continuous inverse. The ideas which we have sketched here will be presented in a more orderly fashion in Chapter II.

1.4 Asymptotic Controllability

As pointed out in the previous section, mappings defined by equations such as

$$\dot{y}(t) + y(t) = \dot{x}(t) - x(t) \tag{1.5}$$

do not have continuous inverses because their inverse equation is unstable. Yet, such systems can be controlled to a large extent. For example, one can find a bounded input x such that y(t) is identical to some desired function over some <u>finite</u> interval $[o, \sigma]$. One may also find x(t) such that y(t) takes on a given value at a

given time or, one may find a bounded $x(t)$ such that $y(t)$ approaches any desired value asymptotically. Systems having this last property are especially important in process control. Thus the fact that a system does not have a continuous inverse does not imply that the system cannot be controlled at all but rather, it merely implies that perfect control is not possible over the entire interval $[0,\infty]$.

Although we will not attempt to define exactly what the capabilities and limitations of a system with an unstable inverse equation are, we will study the conditions under which steady-state control is possible. If a system can be controlled in the steady-state we will call it asymptotically controllable. A precise definition will be given in (2.5). In what follows here, we will try to give the physical motivation for that definition.

Since any reasonable definition of asymptotic controllability requires that the limit as t approaches infinity of $Y(t)$ should exist, it is clear that asymptotic stability of the differential equations is a prerequisite. It is, however, not sufficient to insure that it will be possible to find an X such that $X(t)$ tends toward a preassigned constant. An obvious example is the one-input, two-output system

$$\begin{aligned} \dot{y}_1(t) + y_1(t) &= x(t) \\ \dot{y}_2(t) + 2y_2(t) &= x(t) \end{aligned} \qquad (1.6a)$$

Although these equations are asymptotically stable, it is not possible to find x such that $y_1(t) \longrightarrow a$ and $y_2(t) \longrightarrow b$ as $t \longrightarrow \infty$ except in the special case where $2a = b$.

At the same time, it should not be assumed that there are no one-input, two-output systems which are capable of meeting these requirements. Consider the nonlinear system

$$\dot{y}_1(t) = (x(t))^3 \qquad (1.7a)$$

$$\dot{y}_2(t) = x(t) \qquad (1.7b)$$

A little experimenting with this system convinces one that by manipulating the single input $x(t)$ it is possible to make the outputs assume any two preassigned steady-state values. This behavior is very interesting and seems to be present in most systems which incorporate some means for self-identification or adaption. However, because of its complexity, it will not be discussed here.

From a more formal point of view the problem of determining the conditions under which it is possible to force the outputs to a preassigned steady-state value may be put into our general framework as follows. First observe that we are only interested in the limit of $Y(t)$ as $t \to \infty$ so that any two responses such that $\lim_{t \to \infty} (Y(t) - Y'(t)) = 0$ are equivalent. In fact, this is an equivalence relation in the precise sense[7] and as such it partitions the subset of C for which $\lim_{t \to \infty} Y(t)$ exists into equivalence classes. A suitable norm for this set of equivalence classes is $|\lim_{t \to \infty} Y(t)| = ||\underline{Y}||$ where $\underline{Y}$ represents the equivalence class containing $Y(t)$. It follows that the conditions under which it is possible to achieve steady-state control are the same as those which guarantee that the system maps C onto this set of equivalence classes.

In this section, and those preceding it, we have attempted to provide motivation for the material to be presented in Chapter II. The purpose has been to engender a point of view which may be summarized as follows: Once the initial state is fixed a system can be viewed as a mapping from a function space S_x onto a second set S_y. By properly selecting the norms for these spaces it is possible to relate many of the important problems in control theory to the study of these mappings.

1.5 Background

The problem of determining the limitations which the equations of motion impose on the control of a physical process has been studied from many points of view, usually with the term controllable being applied to systems which have some desired property. Thus the works of Smith[49], Eckman[12], Kalman[23], and Antrosiewicz[3], to name a few, all contain different definitions of controllability.

In his study of multivariable systems, Mesarović[38,40,41] studied and quantified the concept of interaction and our general approach to the problem of determining the conditions under which a multivariable system can produce a desired output has been very much influenced by his work in this area. The work of Kavanaugh[28,29], Amara[1], and Freeman[14,15] should also be mentioned as being in the same spirit as this investigation.

Although we have used a vector differential equation in representing the equations of the system we have also consistently placed in evidence a second equation relating the output to the

states. This is often not done. In fact, one of the effects of the recent trend toward modeling systems by first order vector differential equations has been to focus attention on the state, often to the extent of leaving the outputs undefined. This is evident, for example, in the recent work of Kalman, Ho, and Narandra[27], Gilbert[18], Marcus and Lee[35], and Antrosiewicz[3], where controllability is defined in terms of the state rather than the outputs.

To emphasize the difference between the point of view taken by this last group of authors and the approach taken here, consider the following definition from reference 27: "A system is controllable if any initial state can be transferred to any desired state in a finite length of time by some control action." It has been shown that a necessary and sufficient condition for the equation (1.2a) to be controllable is that the matrix $(B, AB, \dots A^{p-1}B)$* be of rank p where p is the dimension of Z(t). According to such a definition the system defined by (1.6) is controllable. This clearly points out that controllability does not imply that one can select the inputs so as to obtain an arbitrary response. What it does insure is that one can select the inputs so as to obtain an arbitrary value of the state (or output) at some one point in time.

For those cases in which the system equations are of the form of (1.2) and D is zero Kalman[26] has shown that the concept of observability has certain applications. Without attempting a detailed

*See section 2.1 for an explanation of this notation.

explanation, the condition under which a system of the form (1.2) with $D = 0$ is observable is that $(C^T, A^T C^T, \ldots A^{T\,p-1} C^T)$ be of rank p where again p is the dimension of $Z(t)$. Since the system (1.6) can be represented as a controllable and observable system it is clear that it is possible to have systems which are both controllable and observable in the sense of Kalman but which do not correspond to systems which would, for example, make suitable servomechanisms.

Similar comments apply to most of the other definitions of controllability which have been proposed in connection with optimal control problems. The terms "accessibility" and "reachable" which appear in the papers of Hermann[32] and Roxin[32] also fall into this class.

The various properties of the solutions of differential equations which are used in Chapter II may be found in any of the standard texts[10,33,42]. Particular facts which will be required include the usual existence and uniqueness theorems and the fact that if the right hand side has continuous partial derivatives with respect to the state vector then the solutions are continuous functions of the initial state. We also use the fact that the solution of the linear, time-invariant differential equation

$$\dot{Z}(t) = AZ(t) + BX(t) \tag{1.8}$$

can be written as

$$Z(t) = e^{At} Z(0) + e^{At} \int_0^t e^{-As} BX(s)\,ds \tag{1.9}$$

The idea of looking at a nonhomogeneous differential equation as defining a mapping between function spaces is not discussed in the above references. It has been used by Massera and Schaffer[37], for example, in their study of stability, however, not with the same norms or with the same objective. The work in section 2.4 depends strongly on lemma 2 which is closely related to certain results on total stability[19,36]. Our proof of this lemma is based on Lefschetz's proof of Dychman's theorem[33].

1.6 Summary of Results

In Chapter II the conditions under which a linear, time-invariant system such as (1.2) has an inverse is derived. Explicit formulas are given for the inverses in a number of special cases, including the most general single-input, single-output system. One general result concerning n-input, n-output systems of the form of (1.2) is given by theorem 3 which asserts that if $(D, CB, CAB, \ldots CA^{p-1})$ is of rank n then it will be possible to solve (1.2) for X in terms of Y and its derivatives. It should be noted that only in the case where n = 1 is this condition implied by the controllability and observability conditions cited previously.

If the properties derived for linear systems are to be useful it is important to show that small nonlinearities do not disturb them. Some results in this direction are contained in section 2.4 where it is shown that in the neighborhood of zero conclusions about systems of the form

$$\dot{Z}(t) = AZ(t) + BX(t) + BQ(X(t),Z(t)) \qquad (1.10a)$$

$$Y(t) = CZ(t) \qquad (1.10b)$$

can be derived from an examination of the linear part provided $Q(X(t),Z(t))$ satisfies certain smoothness requirements and contains no linear terms.

Insofar as steady-state control is concerned our principle result is that if $CA^{-1}B + D$ singular then the system (1.2) is not asymptotically controllable. The final section of Chapter II interprets the previous results in terms of transfer matrices.

In Chapter III the effects of feedback are analyzed. It is shown that if a system is linear then under mild restrictions linear derivative feedback does not affect irreducibility, nonsingularity, or asymptotic controllability. It is also shown that nonlinear feedback of a certain type does not affect the stability of the inverse equation. The problem of absolute stability is examined and it is shown that stability of the inverse equation is a requirement for absolute stability.

An additional use for the inverse equation is developed in Chapter IV where some nonstandard time-optimal control problems are discussed and solved. It is shown how the structure of the time-optimal input depends on the inverse equation and how the inverse equation can be used to obtain a parametric representation for the optimum input.

CHAPTER II

TRANSFORMATIONS DEFINED BY DIFFERENTIAL EQUATIONS

2.1 Introduction

Our objective is to study a class of transformations defined by nonhomogeneous differential equations. The basic idea is the following. Assume we are given a first order vector differential equation

$$\dot{Z}(t) = F(X(t),Z(t)) \tag{2.1a}$$

and a function $Y(t)$ defined in terms of $Z(t)$ and $X(t)$

$$Y(t) = G(X(t),Z(t)) \tag{2.1b}$$

If for a given value of $Z(0)$, the differential equation has a unique solution for all functions in a given class S_x then the system (2.1) maps this class of functions into a second set S_y. By introducing norms in these function spaces it is possible to study the continuity and invertibility of these mappings in a systematic way. It may be observed that ordinary point-transformations of the type $Y(t) = G(X(t))$ can be thought of as being a special case and are perhaps the simplest examples of the type of transformations we are considering. The motivation for this study can be found in later chapters where applications to various problems in automatic control are discussed.

Lower case letters refer to scalars, upper case to vectors and matrices. If A is an arbitrary vector or matrix then A^T is its transpose, if A is square and nonsingular A^{-1} denotes its inverse.

The $i^{\underline{th}}$ element of a vector X will be written as x_i, the $ij^{\underline{th}}$ element of a matrix A will be written as a_{ij}. Occasionally A will be written as (a_{ij}). If $A_1, A_2, \ldots A_n$ are a set of matrices or vectors all having the same number of rows then $(A_1, A_2, \ldots A_n)$ denotes the matrix consisting of all the columns of $A_1, A_2, \ldots A_n$. Thus, if the A_i are p by p matrices then the matrix $(A_1, A_2, \ldots A_n)$ is p by np. For scalars $|a|$ denotes the absolute values, for vectors and matrices $|A|$ denotes the sum of the absolute values of the elements of A. The determinant of A will be written as det A.

2.2 Function Spaces

Let E_n denote Euclidean n-space and let $\underline{\sigma}$ and I denote the intervals $[0,\sigma]$ and $[0,\infty]$ respectively. The set of all bounded continuous functions with domain $\underline{\sigma}$ and range E_n will be denoted by $C_n(\underline{\sigma})$. The set of bounded functions with domain I and range E_n which are continuous on every finite interval contained in I will be denoted by C_n. For the function space $C_n(\underline{\sigma})$ we use the norm

$$||X||_\sigma = \sup_{t \varepsilon \underline{\sigma}} |X(t)| \qquad (2.2)$$

In the sequel we shall always assume that σ is finite and nonzero without specifying it each time. For C_n the norm is defined in the same way and written as

$$||X|| = \sup_{t \varepsilon I} |X(t)| \qquad (2.3)$$

Let $C_n^k(\underline{\sigma})$ denote the subset of $C_n(\underline{\sigma})$ consisting of those

functions which have at least k continuous derivatives on $\underline{\sigma}$ and let C_n^k be the corresponding subset of C_n. If $X \varepsilon C_n^k(\underline{\sigma})$ or C_n^k we denote its $k^{\underline{th}}$ derivative by $X^{(k)}$. For our purposes it is indispensable to have norms for the spaces $C_n^k(\underline{\sigma})$ and C_n^k which reflect the magnitude X and its first k derivatives. With this in mind we define the $C_n^k(\underline{\sigma})$ norm as

$$||X||_{\sigma}^{k} = \sup_{0 \leq i \leq k} ||X^{(i)}(t)||_{\sigma} \tag{2.4}$$

The norm for C_n^k is written as $||X||^k$ and is given by (2.4) with $\underline{\sigma}$ replaced by I.

To see that (2.4) actually defines a norm notice that i) $||X||_{\sigma}^{k}$ is nonnegative and vanishes if and only if $X(t) \equiv 0$ on $\underline{\sigma}$, ii) $||aX||_{\sigma}^{k} = a \cdot ||X||_{\sigma}^{k}$, and iii) $|X^{(i)}(t)-Y^{(i)}(t)| \leq |X^{(i)}(t)| + |Y^{(i)}(t)|$ for all i and t and hence it follows that the triangle inequality is satisfied. It is manifest that these spaces are linear so that (2.4) makes $C_n^k(\underline{\sigma})$ a normed linear space. When we speak of the continuity of a mapping between two such spaces it is always continuity with respect to the norm topology which we have in mind.

2.3 Linear Time-Invariant Systems

Let X(t) and Y(t) be n-vectors, let Z(t) be a p-vector, and suppose

$$\dot{Z}(t) = AZ(t) + BX(t) \tag{2.5a}$$

$$Y(t) = CZ(t) \tag{2.5b}$$

where A, B, and C are constant. It is well known that if $X \varepsilon C_n(\underline{\sigma})$ then for a given value of Z(0) there exists a unique $Z \varepsilon C_p^1$ satisfying (2.5a) and that Z(t) is given by

$$Z(t) = e^{At} Z(0) + \int_0^t e^{A(t-s)} BX(s)ds \qquad (2.6)$$

Therefore, if Z and Z' are the images of X and X' respectively then $||Z-Z'||_\sigma \leq ||X-X'||_\sigma \cdot \sigma||e^{At}B||_\sigma$. Using this and (2.5a) it follows that $||Z-Z'||_\sigma^1 \leq ((1 + |A|) \cdot \sigma||e^{At}B||_\sigma + |B|)\ ||X-X'||_\sigma$. Since e^{At} is bounded on any finite interval this shows that for finite σ (2.5a) defines a continuous mapping of $C_n(\underline{\sigma})$ into $C_p^1(\underline{\sigma})$.

A sufficient condition for (2.5a) to define a continuous mapping of C_n into C_p^1 is that the eigenvalues of A have negative real parts. This is easily established from the fact that if the eigenvalues of A have negative real parts then there exists α and $\lambda > 0$ such that $||e^{At}B|| \leq \alpha e^{-\lambda t}$. This together with (2.6) gives

$$||Z-Z'|| \leq ||X-X'|| \cdot ||\int_0^\infty \alpha e^{-\lambda(t-s)} ds|| \qquad (2.7)$$

$$\leq ||X-X'|| \alpha/\lambda$$

From this and (2.5a) the continuity of the mapping of C_n into C_p^1 is easily established.

If no additional restrictions are imposed it is possible that for some values of Z(0) (2.5a) may define a continuous mapping of C_n into C_p^1 even though A has eigenvalues with zero or positive real parts. To eliminate this pathological behavior and to prevent similar difficulties from arising when discussing the relationship

between X and Y we now assume: i) That of the np columns in the matrix $(B, AB, \ldots A^{p-1}B)$ there are p linearly independent ones, and ii) That of the np columns in the matrix $(C^T, A^T C^T, \ldots A^{T\,p-1} C^T)$ there are p linearly independent ones.

Systems having this property will be called <u>irreducible</u>. The role of these assumptions has been studied by Gilbert[18] and Kalman[24] and it has been shown that any linear time-invariant system has an irreducible representation. If (2.5b) is replaced by $Y(t) = CZ(t) + DX(t)$ then we will say that the system is irreducible if the corresponding system obtained by setting $D = 0$ is irreducible. For our purpose the importance of irreducibility stems from the following lemma which asserts that if the system is irreducible then asymptotic stability and continuity are equivalent.

<u>Lemma 1</u>: Assume that the system (2.5) is irreducible. Then: i) The differential equation (2.5a) defines a continuous mapping of the input function space C_n into the state function space C_p if and only if the eigenvalues of A have negative real parts. ii) The system (2.5) defines a continuous mapping of the input function space C_n into the output function space C_n if and only if the eigenvalues of A have negative real parts.

<u>Proof</u>: i) The sufficiency of the condition on the eigenvalues has been shown above. To show necessity consider first the mapping of C_n into C_p. From equation (2.6) it follows that the mapping will not be continuous unless $|e^{At}B|$ tends to zero with increasing time. Now suppose A has an eigenvalue with a nonnegative real part. This implies that $|e^{At}|$ does not approach zero as $t \longrightarrow \infty$. More-

over, since $(B,AB,\ldots A^{p-1}B)$ is of rank p the assumption that A has an eigenvalue with a nonnegative real part implies that $|e^{At}(B,AB,\ldots A^{p-1}B)|$ does not approach zero as $t \longrightarrow \infty$. But since A and e^{At} commute this implies that $|(e^{At}B,Ae^{At}B,\ldots A^{p-1}e^{At}B)|$ does not approach zero and hence that $|e^{At}B|$ does not approach zero. ii) Now consider the mapping of X into Y. The sufficiency follows from the fact that Y depends continuously on Z. To prove the necessity notice that since Y(t) is given by

$$Y(t) = Ce^{At}Z(0) + \int_0^t Ce^{A(t-s)}BX(s)ds \qquad (2\text{-}8)$$

it follows that the mapping will not be continuous unless $Ce^{At}B$ approaches zero as $t \longrightarrow \infty$. It remains to prove that this cannot happen if A has a nonnegative eigenvalue. To show that $Ce^{At}B$ does not approach zero if A has a nonnegative eigenvalue we use the fact that $e^{At}B$ does not approach zero and the fact that $(C^T,A^TC^T,\ldots A^{Tp-1}C^T)$ is of rank p. Together these imply that $(C^T,A^TC^T,\ldots A^{Tp-1}C^T)^Te^{At}B$ does not vanish. Using the fact that A and e^{At} commute it follows that $Ce^{At}(B,AB,\ldots A^{p-1}B)$ does not approach zero. Now suppose $Ce^{At}B \longrightarrow 0$ then because $Ce^{At}B$ is a sum of exponential terms it follows that if $Ce^{At}B \longrightarrow 0$ then $d/dt\,Ce^{At}B = CAe^{At}B \longrightarrow 0$. Continuing in this way it is seen that the assumption that $Ce^{At}B \longrightarrow 0$ leads to a contradiction of the statement that $|Ce^{At}(B,AB,\ldots A^{p-1}B)|$ does not approach zero. Q.E.D.

It is also true that irreducibility implies that no linear combination of the $z_j(t)$ vanishes for all possible choices of X(t). To prove this assume the contrary, i.e., assume there exists a

constant, nonzero p-vector F such that $F^TZ(t) \equiv 0$. From (2.5a) it follows by successive differentiation that $F^TB, F^TAB, \ldots F^TA^{p-1}B$ all vanish. This, however, contradicts the assumption that the matrix $(B, AB, \ldots A^{p-1}B)$ contains p linearly independent columns.

In general the mappings discussed in lemma 1 will be continuous with respect to stronger topologies as well. To see this consider the following equations which follow from (2.5) by successive differentiations (assuming momentarily that $X \varepsilon C_n^{p-2}$).

$$\begin{aligned}
CZ(t) &= Y(t) \\
CAZ(t) &= Y^{(1)}(t) - CBX(t) \\
CA^2Z(t) &= Y^{(2)}(t) - CABX(t) - CBX^{(1)}(t) \qquad (2.9) \\
&\ldots\ldots\ldots\ldots\ldots\ldots \\
CA^{p-1}Z(t) &= Y^{(p-1)}(t) - CA^{p-2}BX(t) - CA^{p-3}BX^{(1)}(t) \ldots CBX^{(p-2)}(t)
\end{aligned}$$

If $CA^iB = 0$ for $i = 0, 1, \ldots k-2$ then it follows that $Y^{(k)}(t)$ depends continuously on $Z(t)$ and $X(t)$ and that $Y^{(i)}(t)$ depends continuously on $Z(t)$ for $0 \leq i \leq k$. It follows that if the above condition on CA^iB is satisfied then there exists c such that for finite σ $||Y-Y'||_\sigma^k \leq c||X-X'||_\sigma$. If the eigenvalues of A have negative real parts then this condition on CA^iB implies that there exists a c such that $||Y-Y'||^k \leq c||X-X'||$.

Equation (2.9) also asserts that if $CA^iB = 0$ for $i = 0, 1, \ldots$ k-2 then it will be possible to solve for $X(t)$ in terms of $Y^{(k)}(t)$ and $Z(t)$ if and only if $CA^{n-1}B$ is nonsingular. The if part is obvious and the only if follows by noting that if $CA^{k-1}B$ is singular then any solution for $X(t)$ will necessarily involve higher

derivatives of $Y(t)$. The following theorem subsumes most of the informal statements on continuity given thus far and gives necessary and sufficient conditions for the existence of one type of an inverse mapping both on a finite interval and in the infinite case.

Theorem 1: Consider the system (2.5) where A, B, and C are constant and $X(t)$, $Y(t)$, and $Z(t)$ are n, n, and p-vectors respectively. Assume that the system is irreducible. Then for an arbitrary, fixed, initial value $Z(0)$ it follows that:

i) Equation (2.5) defines a continuous mapping of $C_n(\underline{\sigma})$ into $C_n^k(\underline{\sigma})$ if and only if $CA^iB = 0$ for $i = 0, 1, \ldots k-2$.

ii) If (2.5) maps $C_n(\underline{\sigma})$ into $C_n^k(\underline{\sigma})$ then it has a continuous inverse which maps $C_n^k(\underline{\sigma})$ into $C_n(\underline{\sigma})$ if and only if $CA^{k-1}B$ is nonsingular.

iii) Equation (2.5) defines a continuous mapping of C_n into C_n^k if and only if the eigenvalues of A have negative real parts and $CA^iB = 0$ for $i = 0, 1, \ldots k-2$.

iv) If (2.5) maps C_n into C_n^k then it has a continuous inverse which maps C_n^k into C_n if and only if $CA^{k-1}B$ is nonsingular and $A-B(CA^{k-1}B)^{-1}CA^k$ has p-nk eigenvalues with negative real parts.

Proof: i) From (2.9) it follows that (2.5) maps $C_n(\underline{\sigma})$ into $C_n^k(\underline{\sigma})$ if and only if $CA^iB = 0$ for $i = 0, 1, \ldots k-2$. To see that this mapping is continuous let $X(t)$, $Y(t)$, and $Z(t)$ be one solution of (2.5) and let $X'(t)$, $Y'(t)$, and $Z'(t)$ be a second solution such that $Z(0) = Z'(0)$. Then as shown above $||Z-Z'||_\sigma \leq c||X-X'||_\sigma$ for some c. From (2.9) it follows that $||Y^{(i)}-Y'^{(i)}||_\sigma \leq c_i||Z-Z'||_\sigma$

for i = 0, 1, ... k-1 and that $||Y^{(k)}-Y'^{(k)}||_\sigma \leq c_k||Z-Z'||_\sigma + c'||X-X'||_\sigma$. The result now follows from the definition of the $\mathbb{C}_n^k(\underline{\sigma})$ norm.

ii) If (2.5) maps $\mathbb{C}_n(\underline{\sigma})$ into $\mathbb{C}_n^k(\underline{\sigma})$ then from i) it follows that $\mathbb{C}A^iB = 0$ for i = 0, 1, ... k-2. Therefore $\mathbb{C}A^{k-1}BX(t) = Y^{(k)}(t) - CA^kZ(t)$. If $T^{-1} = CA^{k-1}B$ is nonsingular then (2.5) yields

$$\dot{Z}(t) = (A-BTCA^k)Z(t) + BTY^{(k)}(t) \tag{2.10a}$$

$$X(t) = TY^{(k)}(t) - TCA^kZ(t) \tag{2.10b}$$

From i) it follows that the mapping of $C_n^k(\underline{\sigma})$ into $C_n(\underline{\sigma})$ defined by these equations is continuous. However, if T^{-1} is singular then, as noted above, it will be impossible to solve for X(t) in terms of $Y^{(k)}(t)$ and Z(t) without introducing higher derivatives of Y(t) and hence the inverse will not be defined for some elements of $\mathbb{C}_n^k(\sigma)$.

iii) The necessity of the condition on $\mathbb{C}A^iB$ follows exactly as in i). The necessity of the condition on the eigenvalues follows from the assumed irreducibility and lemma 1. The sufficiency follows as in i) making use of (2.7).

iv) If (2.5) maps $\mathbb{C}_n(\underline{\sigma})$ into $\mathbb{C}_n^k(\underline{\sigma})$ then from iii) it follows that $\mathbb{C}A^iB = 0$ for i = 0, 1, ... k-2. The necessity of $T^{-1} = CA^{k-1}B$ being nonsingular follows as in ii). Assuming this we first establish that the nk rows of the matrices $C, CA, \dots CA^{k-1}$ are all linearly independent. This will be done by showing that no non-trivial linear combination of these rows vanish. Let M_i be arbitrary n x n matrices. From our previous remarks it follows

that $(M_1 C+M_2 CA \ldots M_k CA^{k-1})B = M_k CA^{k-1}B$. Since the right hand side of this equation vanishes if and only if M_k does it follows that no linear combinations of rows involving CA^{k-1} vanishes. Repeating this argument on the equation $(M_1 C+M_2 CA, \ldots M_{k-1} CA^{k-2})AB = M_{k-1} CA^{k-1}B$ we obtain the same conclusion about CA^{k-2} etc.

Let P be any nonsingular p x p matrix whose first n rows are the rows of C, whose 2nd n rows are the rows of CA... and whose $k^{\underline{th}}$ n rows are the rows of CA^{k-1}. Define W(t) as PZ(t) and note that from (2.9) it follows that the first n components of W are the n components of Y, the second n are the n components of $Y^{(1)}$... and the $k^{\underline{th}}$ n components are the components of $Y^{(k-1)}$. With this change of variables the system (2.10) becomes

$$\dot{W}(t) = P(C-BTCA^k)P^{-1}W(t) + PBTY^{(k)}(t) \qquad (2.11a)$$

$$X(t) = -TCA^kP^{-1}W(t) + TY^{(k)}(t) \qquad (2.11b)$$

It is convenient to partition the differential equations (2.11a) into two parts as follows. Let $W_1(t)$ denote the first nk components of W(t) and $W_2(t)$ the remaining p-nk. Let G_1 denote the first nk rows of PBT and let G_2 denote the remaining p-nk. Partition $P(A-BTCA^k)P^{-1}$ into F_{11}, F_{12}, F_{21}, and F_{22} so that (2.11a) becomes

$$\dot{W}_1(t) = F_{11}W_1(t) + F_{12}W_2(t) + G_1Y^{(k)}(t) \qquad (2.12a)$$

$$\dot{W}_2(t) = F_{21}W_1(t) + F_{22}W_2(t) + G_2Y^{(k)}(t) \qquad (2.12b)$$

Since G_1 equals the first nk rows of $PB(CA^{k-1}B)^{-1}$ it follows from the definition of P and the nature of CA^iB that the first

$n(k-1)$ rows of G_1 vanish and the remaining n form an $n \times n$ identity matrix I_n. From the known relationship between $\dot{W}$ and W it follows that $F_{12} = 0$ and F_{11} has the form

$$F_{11} = \begin{bmatrix} 0 & I_n & 0 & \dots & 0 \\ 0 & 0 & I_n & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & I_n \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix} \tag{2.13}$$

Because F_{12} is zero the set of eigenvalues of $P(A-BTCA^k)P^{-1}$ is the union of the sets of eigenvalues of F_{11} and F_{22}. Since $\lambda = 0$ is an eigenvalue of F_{11} of multiplicity nk it follows that the only nonzero eigenvalues of $P(A-BTCA^k)P^{-1}$ are the nonzero eigenvalues of F_{22}.

If $||Y||^k$ is finite then it follows that W_1 is finite as well. From this and (2.12b) it follows that (2.11) defines a continuous mapping of C_n^k into C_n provided that all the eigenvalues of F_{22} have negative real parts. Since the eigenvalues of matrix are unaffected by a similarity transform it follows that all the eigenvalues of F_{22} have negative real parts if and only if $A-BTCA^k$ has p-nk eigenvalues with negative real parts. This shows the sufficiency of the condition on F_{22}.

As a first step in proving the necessity of this condition on the eigenvalues of $A-BTCA^k$ we will show that (2.10) is irreducible. To do this it suffices to show that the ranks of $(BT, (A-BTCA^k)BT, \dots (A-BTCA^k)^{p-1}BT)$ and $(Q, (A-BTCA^k)^TQ, \dots$

$(A-BTCA^k)^{T\,p-1}Q)$ are both p where $Q = (TCA^k)^T$. Consider the first of these. If this matrix is not of rank n then there exists a nonzero p-vector F such that $F^T(BT,(A-BTCA^k)BT,\ldots(A-BTCA^k)^{p-1}BT) = 0$. But this implies that $F^TBT = 0$ and hence that $F^TB = 0$. Also for such an $F, F^T(A-BTCA^k)BT = 0$ and therefore $F^TAB = 0$. Continuing in this way we see that $F^T(B,AB,\ldots A^{p-1}B) = 0$ which contradicts the assumptions that the original system was irreducible.

To show that the second of these matrices is of rank p we show that for no nonzero p-vector F is $F^T(Q,(A-BTCA^k)^TQ,\ldots$ $(A-BTCA^k)^{T\,p-1}Q) = 0$. Suppose it is. Then $TCA^kF = 0$ and consequently $CA^kF = 0$. Since $FTCA^k(A-BTCA^k)F = 0$ it follows that $TCA^k(A-BTCA^k)F = TCAA^kF = 0$ and in general that $CA^iA^kF = 0$. But if this were the case then A^kF would be a nonzero vector such that $(A^kF)^T(C^T,A^TC^T,\ldots A^{T\,p-1}C) = 0$ which contradicts the assumption that the system (2.5) was irreducible.

It is easily verified that if (2.10) is irreducible then (2.11) is as well. It remains only to show that this irreducibility implies that for some choice of Y, regardless of how small $||Y||^k$ is required to be, $Z(t) \longrightarrow \infty$. If F_{22} has any eigenvalues with positive real parts, then it follows that $|e^{F_{22}t}| > \alpha e^{\lambda t}$ for some α and $\lambda > 0$. Consequently, regardless of how small $F_{21}W_1(t) + G_2Y^{(k)}(t)$ is required to be, $Y(t)$ can be chosen so that $W_2(t)$ grows without bound. This leaves the only case where F_{22} has zero eigenvalues. Suppose F_{22} has a zero eigenvalue. Then there exists a (p-nk)-vector Q such that $QF_{22} = 0$ and hence $Q\dot{W}_2(t) = QF_{12}W_1(t) +$

$QG_2Y^{(k)}(t) = Q_oY(t) + Q_1Y^{(1)}(t)\ldots + Q_kY^{(k)}(t)$. If Q_o is nonzero then clearly $QW_2(t)$ grows without bound if $Y(t)$ is chosen such that $Q_oY(t) > 0$. If Q_o is zero then it follows that QW_2 can be expressed in terms of $W_1(t)$ simply by integrating the equation $Q\dot{W}_2(t) = QF_{12}W_1(t) + QG_2Y^{(k)}(t)$. But this contradicts the proven fact that the inverse system is irreducible and completes the proof.

This theorem can be generalized in various ways. One such generalization concerns the system

$$\dot{Z}(t) = AZ(t) + BX(t) \tag{2.14a}$$

$$Y(t) = CZ(t) + DX(t) \tag{2.14b}$$

which is of interest in certain optimal control problems. If D is zero then this is just the system discussed in theorem 1. If D is nonzero the following theorem applies.

Theorem 2: Consider the system (2.14) where A, B, C, and D are constant, D is nonzero, and $X(t)$, $Y(t)$, and $Z(t)$ are n, n, and p-vectors respectively. Assume that the system is irreducible. Then for an arbitrary, fixed, initial value $Z(0)$ it follows that

i) Equation (2.14) defines a continuous mapping of $\mathbb{C}_n(\underline{\sigma})$ into $C_n(\underline{\sigma})$.

ii) The system (2.14) has a continuous inverse which maps $\mathbb{C}_n(\underline{\sigma})$ into $\mathbb{C}_n(\underline{\sigma})$ if and only if D is nonsingular.

iii) The system (2.14) maps $\mathbb{C}_n$ into $\mathbb{C}_n$ if and only if the eigenvalues of A have negative real parts.

iv) If (2.14 maps C_n into C_n then it has a continuous inverse which maps C_n into C_n if and only if all the eigenvalues of $A-CD^{-1}B$

have negative real parts.

<u>Proof</u>: The proof of i) follows from the equation $Y(t) = Ce^{At}Z(0) + \int_0^t Ce^{A(t-s)}BX(s)ds + DX(t)$. If D is nonsingular then the inverse of (2.14) is

$$Z(t) = (A-BD^{-1}C)Z(t) + BD^{-1}Y(t) \tag{2.15a}$$

$$X(t) = -D^{-1}CZ(t) + D^{-1}Y(t) \tag{2.15b}$$

so that if D is nonsingular the inverse exists and from i) it is continuous. If D is singular it follows from the assumption of irreducibility that it is impossible to express X(t) in terms of Z(t) and Y(t) without introducing higher derivatives of Y(t). The proof of iii) and iv) proceed in the same way as the corresponding parts of Theorem 1 provided (2.10) is replaced by (2.15). Q.E.D.

Theorems 1 and 2 give necessary and sufficient conditions for the existence of a certain type of inverse, that is, either one which maps $C_n^k(\underline{\sigma})$ into $\mathbb{C}_n(\underline{\sigma})$ or else one which maps $\mathbb{C}_n^k$ into $\mathbb{C}_n$. A more general problem is that of determining the circumstances under which it is possible to solve for X(t) in terms of Y(t) and any combination of its derivatives. That is, we seek conditions under which (2.5) implies a system of the form

$$\mathring{Z}(t) = A^*Z(t) + B_oY(t) + B_1Y^{(1)}(t)\ldots B_kY^{(k)}(t) \tag{2.16a}$$

$$X(t) = C^*Z(t) + D_oY(t) + D_1Y^{(1)}(t)\ldots D_kY^{(k)}(t) \tag{2.16b}$$

Although we will not obtain a representation for the inverse equations in the general case, the basic question of existence is

answered by the following theorem.

Theorem 3: Consider the system (2.14) where A, B, C, and D are constant (D may be zero) and assume $X(t)$, $Y(t)$, and $Z(t)$ are n, n, and p-vectors respectively. Then a necessary and sufficient condition that it be possible to solve this system for $X(t)$ in terms of $Z(t)$, $Y(t)$, $Y^{(1)}(t)\ldots$ is that $(D, CB, CAB, \ldots CA^{p-1}B)$ be of rank n.

Proof: The sufficiency follows immediately from the following equations which are a consequence of (2.14).

$$
\begin{aligned}
Y(t) &= CZ(t) + DX(t) \\
Y^{(1)}(t) &= CAZ(t) + CBX(t) + DX^{(1)}(t) \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
Y^{(p)}(t) &= CA^{p}Z(t) + CA^{p-1}BX(t) + CA^{(p-2)}BX^{(1)}(t)\ldots DX^{(p)}(t)
\end{aligned}
\qquad (2.17)
$$

This set is exhaustive in that any relationship between $X(t)$, $Z(t)$ and $Y(t)$ and its derivatives which is implied by (2.14) can be derived from (2.17) without using (2.14). If $(D, CB, CAB, \ldots CA^{p-1}B)$ is of rank n then (2.17) provides n equations relating $X(t)$ and $Z(t)$ and $Y(t)$ and its derivatives. If it is not of rank n then clearly it is impossible to solve for $X(t)$ from (2.17) and hence (2.14). Q.E.D.

This proof also shows that if $(D, CB, CAB, \ldots CA^{p-1}B)$ is of rank n then any p-times differentiable output Y defined on a finite interval can be produced by a bounded continuous X. Thus the relationship between X and Y defined by such a system is linear and one to one.

Notice that it was not necessary to assume that the system was irreducible. If $(D, CB, CAB, \ldots CA^{p-1}B)$ is of rank n we will say that the system is <u>nonsingular</u>. As will be discussed in section (2.6) this terminology is in harmony with the usual definition of non-singularity as applied to transfer matrices.

2.4 Nonlinear Transformations

We now examine the question of how the previous results are affected by the presence of a nonlinear term. Our objective is to show that if inputs are sufficiently small and if the initial state is sufficiently close to the equilibrium point then under some circumstances one can determine the properties of the non-linear system by an examination of the linear terms.

Let N(a) denote the set of elements of $C_n(I)$ whose norms are less than a, let M(b) denote the point set $\{Z(t) \mid |Z(t)| < b\}$, and let M(a,b) denote the point set $\{X(t) \mid |X(t)| < a\} \times M(b)$. Consider the system

$$\dot{Z}(t) = AZ(t) + BX(t) + Q(X(t), Z(t)) \tag{2.18a}$$

$$Y(t) = CZ(t) \tag{2.18b}$$

where Q and its partial derivatives with respect to the components of X(t) and Z(t) are continuous in M(a,b) and vanish at X(t) = 0, Z(t) = 0. It follows from the classical existence and uniqueness theorems[10] that for any given $X \epsilon N(a)$ and any given $Z(0) \epsilon M(b)$ there exists a unique solution of (2.15a) defined on an interval $\underline{\sigma}$ and that $\underline{\sigma}$ may be chosen so small that $Z(t) \epsilon M(b)$ for all $t \epsilon \underline{\sigma}$. By

restricting $Z(0)$ still further, say to $M(b/2)$, it is possible to find an interval $\underline{\sigma}'$ such that any $X \epsilon N(a)$ gives rise to a Z which is defined on $\underline{\sigma}'$ having the property that $Z(t) \epsilon M(b)$ for $t \epsilon \underline{\sigma}'$.

Let $\underline{\sigma}$ denote an interval having the properties of $\underline{\sigma}'$ above. Then (2.18a) defines a mapping of $N(a) \times M(b/2)$ into $C^1_p(\underline{\sigma})$. That this mapping is continuous in $Z(0)$ for fixed X has long been known. To see that it is continuous in X for fixed $Z(0)$ let Z and Z' be the images of X and X' respectively and suppose that $Z(0) = Z'(0)$. It follows from (2.18a) that

$$Z(t)-Z'(t) = e^{At}\int_0^t e^{-As}(BX(s)-BX'(s) + Q(X(s),Z(s))-Q(X'(s),Z'(s))ds \tag{2.19}$$

From the assumptions on the continuity of the partial derivatives of $Q(X(t),Z(t))$ and the fact that the matrix expotential is bounded on any finite interval it follows that for $t \epsilon \underline{\sigma}$ and some k and k'

$$|Z(t)-Z'(t)| \leq k\int_0^t |Z(s)-Z'(s)|ds + k'||X-X'|| \tag{2.20}$$

Now let $v(t) = \int_0^t |Z(s)-Z'(s)|ds$ so that (2.20) becomes $\dot{v}(t)-kv(t) \leq k'||X-X'||$. By treating this inequality in the standard way (Bellman[6] page 35) it is easy to show that $|Z(t)-Z'(t)| < k''||X-X'||$ for some k'' and hence to establish the continuity.

A second preliminary result which is more subtle concerns the conditions under which (2.18a) defines a continuous mapping of $N(a') \times M(b')$ into C^1_p. The result which we establish here is

similar to Hahn's statement of the total stability theorem[19], however, our method of proof is quite different.

Lemma 2: If the eigenvalues of A have negative real parts then there exists a' and b', both positive, such that (2.18a) defines a mapping of $N(a') \times M(b')$ into $\mathbb{C}_p^1$ which is continuous in X for fixed Z(0).

Proof: First we show that there exists a' and b' such that (2.18a) maps $N(a') \times M(b')$ into $\mathbb{C}_p^1$. Let λ be a positive number such that -2λ is greater than the real part of any of the eigenvalues of A. Select α such that $|e^{At}| \leq \alpha e^{-2\lambda t}$ for all t. Note that since the partial derivatives of Q(X(t),Z(t)) with respect to X(t) and Z(t) are continuous and vanish at X(t) = 0, Z(t) = 0 there exists c and c' such that in M(c,c').

$$Q(X(t),Z(t)) \leq (|X(t)| + |Z(t)|)\lambda/\alpha \tag{2.21}$$

Using this and the above bound on e^{At} it follows from (2.18a) that for some k and all $t \epsilon \underline{\sigma}$

$$|Z(t)| \leq \alpha e^{-2\lambda t}|Z(0)| + \lambda e^{-2\lambda t}\int_0^t e^{2\lambda s}|Z(s)|ds + k\|X\| \tag{2.22}$$

provided $X \epsilon N(c)$ and $|Z(t)| < c'$ for $t \epsilon \underline{\sigma}$. Let $v(t) = \int_0^t e^{2\lambda s}|Z(s)|ds$. Multiply (2.22) by $e^{2\lambda t}$ to get

$$\dot{v}(t) - \lambda v(t) \leq |Z(0)|\alpha + k\|X\|e^{+2\lambda t} \tag{2.23}$$

Now multiply this equation by $e^{+\lambda t}$ and integrate by parts to get

$$v(t)e^{+\lambda t} \leq \left[|Z(0)|\alpha/\lambda + k\|X\|e^{2\lambda t}/\lambda\right]e^{\lambda t} \tag{2.24}$$

This and (2.23) can be seen to imply

$$|Z(t)| \leq 2|Z(0)|\alpha e^{-2\lambda t} + ||X||2k \qquad (2.25)$$

Again, this equation is valid for all $t\epsilon\underline{\sigma}$ provided $|Z(t)| < c'$ for $t\epsilon\underline{\sigma}$. But (2.25) implies that if $||X||$ and $|Z(0)|$ are taken to be sufficiently small then $Z(t)$ can be made less than c' for all t and hence there does in fact exist a' and b' such that (2.18a) maps $N(a') \times M(b')$ into C_p.

To show continuity choose c and c' such that for any X and X' belonging to N(c) and any Z(t) and Z'(t) belonging to M(c') we have

$$Q(X(t),Z(t))-Q(X'(t),Z'(t)) \leq (|X(t)-X'(t)| + |Z(t)-Z'(t)|)\lambda/\alpha \qquad (2.26)$$

Use this in (2.19) to show that for $X\epsilon N(c)$ and some k

$$|Z(t)-Z'(t)| \leq \lambda e^{-2\lambda t}\int_0^t e^{2\lambda t}|Z(s)|ds + k||X-X'|| \qquad (2.27)$$

provided $|Z(t)| < c'$. By treating this in the same way as (2.20) was treated one can show that

$$||Z-Z'|| \leq ||X|| \cdot 2k \qquad (2.28)$$

provided $||X||$ and $|Z(0)|$ are sufficiently small. Q.E.D.

In general the inverse of the system (2.18) will not have a unique solution for a given Y. For this reason we now restrict ourselves to systems of the form

$$Z(t) = AZ(t) + BX(t) + BQ(X(t),Z(t)) \tag{2.29a}$$

$$Y(t) = CZ(t) \tag{2.29b}$$

where $Q(X(t),Z(t))$ satisfies the same conditions as before in $M(a,b)$. By differentiating (2.29b) in the same way as we did (2.5b) previously one obtains

$$\begin{aligned} CZ(t) &= Y(t) \\ CAZ(t) &= Y^{(1)}(t) - CB(X(t) + Q) \qquad (2.30) \\ &\cdots\cdots\cdots\cdots \\ CA^{p-1}Z(t) &= Y^{(p-1)}(t) - CA^{p-2}B(X(t) + Q)\ldots CB(X(t) + Q)^{(p-2)} \end{aligned}$$

From this it can be seen that if it is possible to solve these equations for $Y^{(i)}$ in terms of X and Z the solutions will not involve derivatives of X. In general this is not true for the system (2.18) and this is our reason for restricting our attention to the system (2.29).

The following theorem gives one set of conditions which are sufficient to insure that (2.29) defines a continuous, invertible mapping. It should be observed that since we are interested only in sufficient conditions no assumptions on irreducibility need be made.

<u>Theorem 4</u>: Consider the system (2.29) with the given assumptions on $Q(X(t),Z(t))$ valid in $M(a,b)$. Assume further that $CA^iB = 0$ for $i = 0, 1, \ldots k-2$ then:

i) There exists an interval $\underline{\sigma}$ such that for fixed $Z(0)\epsilon M(b/2)$ this system defines a continuous mapping of $N(a)$ into $C_n^k(\underline{\sigma})$. If

$CA^{k-1}B$ is nonsingular then there exists $c' > 0$ such that for any given $Y \epsilon C_n^k(\underline{\sigma})$ with $||Y||_\sigma^k < c'$ there is a unique $X \epsilon N(a)$ satisfying the system.

ii) If all the eigenvalues of A have negative real parts then there exists a' and b', both positive, such that this system defines a mapping of $N(a') \times M(b')$ into C_n^k which is continuous in X for fixed Z(0). If in addition $CA^{k-1}B$ is nonsingular and if $A-B(CA^{k-1}B)^{-1}CA^k$ has p-nk eigenvalues with negative real parts then there exists c' and d', both positive, such that if $||Y||^k < c'$ and $Z(0) \epsilon M(d')$ then corresponding to Y and Z(0) there is a unique $X \epsilon N(a)$ which satisfies the system.

Proof: i) Since (2.29a) is a special case of (2.18a) it follows that the remarks about existence and uniqueness apply. Equation (2.30) shows that if $CA^iB = 0$ for $i = 0, 1, \ldots k-2$ then $Y^{(i)}(t)$ depends continuously on Z(t) for $i \leq k-1$, and that $Y^{(k)}(t)$ is a continuous function of Z(t) and X(t). From this and the definition of the $C_n^k(\underline{\sigma})$ norm it follows that (2.29) defines a continuous mapping of $C_n(\underline{\sigma})$ into $C_n^k(\underline{\sigma})$ under the given assumptions.

Let $T^{-1} = CA^{k-1}B$. If T^{-1} is nonsingular then it follows from (2.30) and the assumptions on CA^iB that

$$TCA^kZ(t) = TY^{(k)}(t) - X(t) - Q(X(t),Z(t)) \qquad (2.31)$$

Applying the implicit function theorem[17] to this equation gives

$$X(t) = -TCA^kZ(t) + TY^{(k)}(t) + Q'(Z(t),Y^{(k)}(t)) \qquad (2.32)$$

where $Q'(Z(t),Y^{(k)}(t))$ enjoys the same properties as $Q(X(t),Z(t))$ in some set $M(a',b')$. Using this and (2.29) it is seen that the inverse for (2.29) is

$$\dot{Z}(t) = (A-BTCA^k)Z(t) + BTCA^kY^{(k)}(t) + BQ'(Y^{(k)}(t),Z(t)) \tag{2.33a}$$

$$X(t) = -TCA^kZ(t) + TY^{(k)}(t) + Q'(Z(t),Y^{(k)}(t) \tag{2.33b}$$

From the remarks made about (2.18a) it follows that (2.33a) defines a continuous mapping for small X, $Z(0)$, and $\underline{\sigma}$ and hence the mapping of $C_n^k(\underline{\sigma})$ into $C_n(\underline{\sigma})$ defined by 2.33 is continuous.

ii) In view of lemma 2 it remains only to show that there exists c' and d' having the desired properties. Consider (2.33a). Define P from (2.30) exactly as P was defined in the proof of Theorem 1, part iv. Let $W(t) = PZ(t)$. Partitioning W as in that proof we see that $W_1(t)$ is small if $||Y||^k$ is and that $W_2(t)$ satisfies the equation

$$\dot{W}_2(t) = F_{22}W_2(t) + F_{21}W_1(t) + G_2Q'(PW(t),Y^{(k)}(t)) \tag{2.34}$$

If $A-B(CA^{k-1}B)^{-1}CA^k$ has $p-nk$ eigenvalues with negative real parts then because $F_{12} = 0$ all the eigenvalues of F_{22} have negative real parts. The desired result now follows from lemma 2. Q.E.D.

2.5 Asymptotic Controllability

Adequate control is often possible even if a system fails to have a continuous inverse. The reason being that for many processes the exact form of the response is unimportant as long as the proper

steady-state value is eventually achieved. For this reason the problem of determining the circumstances under which steady-state control is possible is of interest. We begin by showing how this problem can be stated in terms of a mapping.

In order to state this problem in terms of a mapping it is necessary to use the notion of an equivalence class. Let $\underline{C}_n$ denote the subset of C_n for which $\lim_{t\to\infty} Y(t)$ exists. The norm for $\underline{C}_n$ is taken to be the usual C_n norm. Partition $\underline{C}_n$ into equivalence classes according to the relationship $Y \sim X$ if $\lim_{t\to\infty} |Y(t)-X(t)| = 0$ and let underlined capitals such as $\underline{Y}$ and $\underline{X}$ denote particular equivalence classes. The set of all equivalence classes will be denoted by R_n and the norm for R_n is defined as $||\underline{Y}|| = \lim_{t\to\infty} |Y(t)|$. We will say that a system is asymptotically controllable if it defines a continuous mapping of $\underline{C}_n$ onto R_n. From a practical point of view this says that if a system is asymptotically controllable then its outputs may be made to approach any desired steady-state value by using an input which tends toward a constant. Moreover, it implies that small changes in the input result in small changes in the asymptotic value of the output. The conditions under which the system (2.14) is asymptotically controllable are given by the following theorem.

Theorem 5: Consider the system (2.14) with A, B, C, and D constant. Let X(t) and Y(t) be n-vectors and assume the system is irreducible. Then these equations define an asymptotically controllable system if and only if the eigenvalues of A have negative

real parts and $CA^{-1}B+D$ is nonsingular.

Proof: The necessity of the condition on the eigenvalues of A follows from lemma 1 and the fact that the mapping is required to be continuous. Assuming that the eigenvalues of A have negative real parts it follows that if $X \epsilon C_n$ then

$$0 = A\underline{Z} + B\underline{X} \tag{2.35a}$$

$$\underline{Y} = C\underline{Z} + D\underline{X} \tag{2.35b}$$

As explained above, the underlines indicate asymptotic values. Since A has no zero eigenvalues it is nonsingular and hence $\underline{Y} = (CA^{-1}B + D)\underline{X}$. Therefore, the nonsingularity of the matrix $CA^{-1}B + D$ is necessary and sufficient. Q.E.D.

It is of some interest to note that the nonsingularity of $(CA^{-1}B + D)$ implies that the matrix $(D, CB, CAB, \ldots CA^{p-1}B)$ appearing in theorem 3 is of rank n but that this implication may not be reversed. The problem of establishing the conditions under which (2.18) is asymptotically controllable in a neighborhood of the origin immediately suggests itself. In view of lemma 2 such a result would seem reasonable, however, a necessary step is to show that under suitable restrictions of $Z(0)$ and $||X||$ (2.18) maps $\underline{C}_n$ into $\underline{C}_n$. That is, one must show that $\dot{Y}(t) \to 0$. If additional restrictions are placed on X, e.g., if it is required that $\int_0^\infty |\dot{X}(t)| dt$ be small, then this can be shown, but no proof indicating that $\dot{Y}(t) \to 0$ for all $X \epsilon \underline{C}_n$ has been found.

2.6 Transfer Matrix Representation

In view of the large body of control systems theory which is based on the use of transfer functions and transfer matrices it is appropriate to interpret our results on linear systems in these terms. Let $\overline{X}(s)$ and $\overline{Y}(s)$ denote the Laplace transforms of $X(t)$ and $Y(t)$ respectively. Then if $X(t)$ and $Y(t)$ are related by (2.14) we have

$$\overline{Y}(s) = (C(Is-A)^{-1}B + D)\,\overline{X}(s) + C(Is-A)^{-1}Z(0) \qquad (2.36)$$

The n x n matrix $C(Is-A)^{-1}B + D$ is usually called the transfer matrix. It is an important property of the system and many of our results which relate to linear systems can be interpreted in terms of it.

Let $P(s) = C(Is-A)^{-1}B + D$. Our objective is to discuss the various assumptions which have been made about A, B, C, and D in terms of $P(s)$. First notice that for all values of s such that $(Is-A)$ is nonsingular $(Is-A)^{-1}$ is bounded and thus the poles of $P(s)$ occur where $\det(Is-A) = 0$. Since the values of s which satisfy this equation are just the eigenvalues of A it follows that if A has eigenvalues with negative real parts then the poles of $P(s)$ lie in the left half-plane. Notice that $\det A = 0$ then $P(s)$ has a pole at the origin.

The number of times that an output is differentiable, i.e. the largest number k such that $Y \epsilon C_n^k$ for all $X \epsilon C_n$, can also be easily related to $P(s)$. Suppose $X(t)$ is such that $\dot{X}(0^+) - \dot{X}(0) = J \neq 0$.

2.6 Transfer Matrix Representation

In view of the large body of control systems theory which is based on the use of transfer functions and transfer matrices it is appropriate to interpret our results on linear systems in these terms. Let $\overline{X}(s)$ and $\overline{Y}(s)$ denote the Laplace transforms of $X(t)$ and $Y(t)$ respectively. Then if $X(t)$ and $Y(t)$ are related by (2.14) we have

$$\overline{Y}(s) = (C(Is-A)^{-1}B + D)\,\overline{X}(s) + C(Is-A)^{-1}Z(0) \tag{2.36}$$

The n x n matrix $C(Is-A)^{-1}B + D$ is usually called the transfer matrix. It is an important property of the system and many of our results which relate to linear systems can be interpreted in terms of it.

Let $P(s) = C(Is-A)^{-1}B + D$. Our objective is to discuss the various assumptions which have been made about A, B, C, and D in terms of $P(s)$. First notice that for all values of s such that $(Is-A)$ is nonsingular $(Is-A)^{-1}$ is bounded and thus the poles of $P(s)$ occur where $\det(Is-A) = 0$. Since the values of s which satisfy this equation are just the eigenvalues of A it follows that if A has eigenvalues with negative real parts then the poles of $P(s)$ lie in the left half-plane. Notice that $\det A = 0$ then $P(s)$ has a pole at the origin.

The number of times that an output is differentiable, i.e. the largest number k such that $Y \epsilon C_n^k$ for all $X \epsilon C_n$, can also be easily related to $P(s)$. Suppose $X(t)$ is such that $\dot{X}(0^+) - \dot{X}(0) = J \neq 0$.

Then by applying the initial value theorem it follows that $\dot{Y}(0^+) - \dot{Y}(0) = \lim_{s \to \infty} P(s)J$. Therefore if $P(\infty) = 0$ then the mapping is into C_n^1. In general the mapping will be into C_n^k if and only if

$$\lim_{s \to \infty} s^i P(s) = 0 \quad \text{for } i = 0, 1, \ldots k-1 \tag{2.37}$$

Since det $P(s)$ is a rational function it either vanishes identically or else is nonzero at all but a finite number of points in the s-plane. In the later case it is possible to solve the equation $\overline{Y}(s) = P(s)\overline{X}(s)$ for $\overline{X}(s)$ and $P(s)$ is said to be nonsingular. The condition that $(D, CB, CAB, \ldots CA^{p-1}B)$ be of rank n implies that $P(s)$ is nonsingular and therefore justifies our terminology.

If the system is nonsingular then $\overline{X}(s) = P^{-1}(s)\overline{Y}(s)$. Since $P^{-1}(s)$ contains poles where det $P(s)$ vanishes it follows that $\overline{X}(s)$ will contain right half-plane poles if det $P(s)$ contains right half-plane zeros. Moreover, the zeros of det $P(s)$ coincide with the eigenvalues of the inverse equation. Much of the utility of the inverse equation stems from this fact. Systems for which det $P(s)$ has no right half-plane zeros are sometimes called minimum phase systems. Thus if the inverse equation is stable the system is minimum phase.

The final value theorem states that if $P(s)$ has no poles in the right half-plane or on the $j\omega$-axis then $\underline{Y} = P(0)\underline{X}$, but $P(0)$ is simply $CA^{-1}B + D$ so that the condition for asymptotic con-

trollability can be expressed in terms of $P(s)$ by saying $P(0)$ should be nonsingular.

A detailed explanation of the implications of irreducibility has been given by Gilbert[18] and Kalman[24]. Roughly speaking irreducibility implies that there are no redundant variables in the state vector. If X and Y are scalars then irreducibility insures that the numerator and denominator of $P(s)$ contain no common factors and that the degree of the denominator is the same as the dimension of Z. Also in the scalar case it may be shown that the degree of the numerator of $P(s)$ is $p-1-k$ where p is the dimension of Z and k is the least nonnegative integer i for which CA^iB is nonzero. By applying the results of theorem 1 we see that if the degree of the denominator of the transfer function exceeds that of the numerator by k and if the denominator contains left half-plane zeros only then the system maps C_1 into C_1^k. The condition for asymptotic controllability is that the numerator polynomial should not have a zero at the origin.

2.7 Conclusions

Our objective has been to develop a means for characterizing nonhomogeneous differential equations which places in evidence certain properties which are important in automatic control applications. It has been shown that the systems under consideration can be viewed as a mapping whose domain is a set of functions S_x and whose range can be either a set of functions or a set of equivalence classes. Moreover, it is possible to introduce norms

for the input and output spaces in such a way as to make it possible to express many of the standard problems in control in terms of the continuity and invertability of these mappings. The idea of computing an inverse equation is introduced and the role of this equation in determining the character of the mapping was discussed.

These ideas have been applied to linear, time-invariant systems and to a certain class of nonlinear systems. The principle results are the following:

1) If the system is linear, time-invariant, and irreducible, then continuity and asymptotic stability are equivalent.

2) If the system is described by (2.5) and $CA^{i}B = 0$ for $i = 0, 1, \ldots k-2$ with $CA^{k-1}B$ being of full rank then explicit forulas can be given for the inverse equation and Theorem 1 provides a complete description of the mapping. The most general single-input single-output system is of this type.

3) The mapping defined by (2.14) has an inverse if and only if the matrix $(D, CB, CAB, \ldots CA^{p-1}B)$ is of rank n. The term nonsingular was used to describe such systems.

4) If certain continuity requirements are satisfied the effect of a nonlinear term of the form $BQ(X(t), Z(t))$ can be ignored if one is only interested in local behavior.

5) The concept of asymptotic controllability was introduced and it was shown that the system (2.14) is asymptotically controllable if and only if the system is asymptotically stable and $CA^{-1}B + D$ is of full rank.

Important problems which remain include that of obtaining a representation of the inverse equation which would apply to the general case and proving that the conditions for asymptotic controllability are not affected by a nonlinear term of the form discussed in theorem 4.

CHAPTER III

THE EFFECT OF FEEDBACK

3.1 Types of Feedback

If the relationships between a systems inputs and outputs are unsatisfactory one often seeks to improve them by the use of feedback. For example, if the system equations are

$$\dot{Z}(t) = F(Z(t), X(t)) \tag{3.1a}$$

$$Y(t) = G(Z(t), X(t)) \tag{3.1b}$$

and if the relationship between X and Y that these equations impose is unsuitable then it may be possible to remedy the situation by forcing $X(t)$ to depend on Y. Ordinarily this has the effect of replacing $X(t)$ in (3.1a) by $X(t) + X'(t)$ where $X'(t)$ is dependent on Y. Much of the work in automatic control theory is concerned with the evaluation of the effects of such substitution.

This chapter is concerned with the question of how feedback affects the properties which were discussed in the previous chapter and is also concerned with its effect on stability. The objective is to define some of its possibilities and limitations. The question of how certain types of feedback affect the properties of irreducibility, stability, and stability of the inverse equation will be examined. Our principle results relate to linear, time-invariant systems with linear and nonlinear feedback.

The dependence of $X'(t)$ on Y can take many forms. If $X'(t)$ is a function of $Y(t)$ only, i.e. $X'(t) = G(Y(t))$ then one speaks of gain feedback. A more general case is where $X'(t) =$

$G(Y(t), Y^{(1)}(t), \ldots Y^{(k)}(t))$. This is usually called derivative feedback and includes gain feedback as a special case. A third possibility is that the relationship between $X'(t)$ and Y is itself a dynamic one being defined by a system of the form

$$\dot{Z}'(t) = F'(Z'(t), Y(t)) \tag{3.2a}$$

$$X'(t) = G'(Z'(t), Y(t)) \tag{3.2b}$$

This is a generalization of what is usually called integral feedback.

Notice that derivative feedback does not change the number of initial conditions needed to describe the motion; that is, it does not change the dimension of the differential equation of the system. This is not true for integral feedback, however, as can be seen by eliminating $X'(t)$ from (3.1) and (3.2). If this is done the resulting system is

$$\dot{Z}(t) = F(Z(t), X(t) + G'(Z'(t), G(Z(t), X(t))) \tag{3.3a}$$

$$\dot{Z}'(t) = F'(Z'(t), G(Z(t), X(t))) \tag{3.3b}$$

$$Y(t) = G(Z(t), X(t)) \tag{3.3c}$$

The system (3.3) clearly has a higher dimensional state vector than the original system. The equations (3.3a) and (3.3b) which determine the stability of the closed loop system constitute what Lefschetz[33] terms "differential equations on product spaces". It is interesting to note the immediate application which the theory of such equations has in automatic control.

In studying the effects and limitations of feedback as a device for altering a given system we seek to identify those properties of the system which are not affected by a class of feedbacks. In general stability is not such a property for it is easy to give examples of both unstable systems which can be stabilized by feedback and stable systems which have been unstabilized by feedback. Other properties such as invertibility, irreducibility, and stability of the inverse equations do, however, remain unchanged by the addition of wide class of feedbacks. Under some circumstances even stability remains unaffected by the addition of feedback as the recent work on the absolute stability problem shows.

The main results given here relate to linear systems having linear and nonlinear derivative feedback. The corresponding problems involving integral feedback seem to be more difficult and are not treated here. It is important to note that in our definition of feedback we have not allowed the feedback to be a function of the entire state vector Z, but instead it is assumed that the feedback relates to the actual output of the system Y. This assumption seems to be the natural one from a physical standpoint although other authors have taken a different point of view.

3.2 Feedback Invariants

As mentioned above the results to be discussed relate to linear, time-invariant systems with linear and nonlinear feedback. Consider the system

$$\dot{Z}(t) = AZ(t) + BX(t) \tag{3.4a}$$

$$Y(t) = CZ(t) \tag{3.4b}$$

to which will be added either linear feedback of the type

$$X'(t) = G_o Y(t) + G_1 Y^{(1)}(t) + \dots G_k Y^{(k)}(t) \tag{3.5}$$

or else nonlinear feedback of the type

$$X'(t) = G(Y(t), Y^{(1)}(t), \dots Y^{(k)}(t)) \tag{3.6}$$

In either case we will say that the feedback is of order k if the $k^{\underline{th}}$ derivative of $Y(t)$ is the highest one appearing in the expression for $X'(t)$.

In view of the role played by irreducibility in the previous chapter the following theorem showing that linear derivative feedback does not affect the irreducibility is of interest.

Theorem 6: Assume that the system (3.4) is irreducible and maps C_n into C_n^k. Then if linear derivative feedback of order k-1 or less is applied the resulting system is still irreducible.

Proof: Since the original system is assumed to map $\mathbb{C}_n$ into $\mathbb{C}_n^k$ it follows from theorem 1 that CA^iB is zero for $i = 0, 1, \dots k-2$, and hence that $Y^{(i)}(t) = CA^iZ(t)$ for the same values of i. The differential equations after feedback can be expressed in terms of Y as

$$\dot{Z}(t) = AZ(t) + BX(t) + B(G_o Y(t) + G_1 Y^{(1)}(t) + \dots G_{k-1} Y^{(k-1)}(t)) \tag{3.7}$$

Expressing $Y^{(i)}$ in terms of $Z(t)$ enables one to write the equations of the feedback system as

$$\dot{Z}(t) = FZ(t) + BX(t) \tag{3.8a}$$

$$Y(t) = CZ(t) \tag{3.8b}$$

where

$$F = A + BG_o C + BG_1 CA + \dots BG_{k-1} CA^{k-1} \tag{3.9}$$

In showing the irreducibility of (3.8) we will use the fact that $CF^i = CA^i$ for $i = 0, 1, \dots k-1$. This follows immediately from the form of F and the fact that $CA^iB = 0$ for $i = 0, 1, \dots k-2$. To show that (3.8) is irreducible it is necessary and sufficient to show that $(B, FB, \dots F^{p-1}B)$ and $(C^T, F^TC^T, \dots F^{Tp-1}C^T)$ are both of rank p where p is the dimension of Z(t). Suppose $(B, FB, \dots F^{p-1}B)$ is not of rank p. Then there exists a nonzero, p-dimensional, row-vector H such that $H(B, FB, \dots F^{p-1}B) = 0$. This implies $HF^iB = 0$ for $i = 0, 1, \dots p-1$. Using the fact that $HB = 0$ it follows that $HAB = HFB$ and hence vanishes. Reasoning inductively, it follows that if $HF^iB = 0$ for $i = 0, 1, \dots j-1$, then $HA^jB = HF^jB$. What this shows is that if $H(B, FB, \dots F^{p-1}B)$ vanishes then $H(B, AB, \dots A^{p-1}B)$ vanishes. This last statement establishes a contradiction because H was assumed to be nonzero and the assumed irreducibility of (3.4) implies that $(B, AB, \dots A^{p-1}B)$ is of rank p.

To show that $(C^T, F^TC^T, \dots F^{Tp-1}B)$ is of rank p again assume the contrary and let H be a nonzero p-dimensional row-vector such that $H(C^T, F^TC^T, \dots F^{Tp-1}B) = 0$. As noted above $CF^i = CA^i$ for $i = 0, 1, \dots k-1$ so the assumption of H implies that $H(C^T, A^TC^T, \dots A^{Tk-1}B) = 0$. But from the form of F it follows that if $CA^{k-1}H^T = 0$ then $FH^T = AH^T$

and hence $CF^kH^T = CA^{k-1}FH^T = CA^kH^T$. Therefore the vanishing of CF^kH^T implies $H(C^T, A^TC^T, \ldots A^{Tk}C^T) = 0$. This argument can now be repeated p-l-k times to show that $H(C^T, A^TC^T, \ldots A^{Tp-1}C^T) = 0$. The last statement contradicts the assumption that (3.4) was irreducible and completes the proof.

The assumption on the order of the feedback was used in an essential way and it is possible to construct examples which show that theorem 6 is not true if the derivative feedback is allowed to be of order k. A second result relating to this type of feedback is the following theorem which shows that with the above assumptions feedback cannot make a nonsingular system singular.

Theorem 7: Assume that the system (3.4) is irreducible and nonsingular and that it maps C_n into C_n^k. Then if linear derivative feedback of order k-l or less is applied the resulting system is also nonsingular and, if the eigenvalues of F have negative real parts then it defines a continuous mapping of C_n into C_n^k.

Proof: To show that the nonsingularity of (3.4) implies the nonsingularity of (3.8) it is necessary to show that if $(CB, CAB, \ldots CA^{p-1}B)$ if the rank n then so is $(CB, CFB, \ldots CF^{p-1}B)$. Assume that the matrix $(CB, CFB, \ldots CF^{p-1}B)$ is not of rank n. Then there exists a nonzero n-vector H such that $H(CB, CFB, \ldots CF^{p-1}B) = 0$. Since our assumptions insure that $CF^i = CA^i$ for i = 0, 1, ... k-l it follows that $H(CB, CAB, \ldots CA^{k-1}B) = 0$. Consider HCF^kB which may be written as $HCA^{k-1}FB = HCA^kB$. This shows $H(CB, CAB, \ldots CA^kB) = 0$. This argument can now be repeated p-l-k times to show that $H(CB, CAB, \ldots$

$CA^{p-1}B) = 0$. This is a contradiction, however, because H was assumed to be nonzero and $(CB, CAB, \ldots CA^{p-1}B)$ is of rank n.

Showing that if the eigenvalues of F have negative real parts then the system with feedback defines a continuous mapping of $\mathbb{C}_n$ into $\mathbb{C}_n^k$ is equivalent by theorem 1 to showing that $CF^iB = 0$ for $i = 0, 1, \ldots k-2$. But as noted above $CF^i = CA^i$ for these values of if and therefore since $CA^iB = 0$ for $i = 0, 1, \ldots k-2$ the desired result is obtained. Q.E.D.

In addition to irreducibility and nonsingularity, asymptotic controllability is also unaltered by the type of linear derivative feedback under discussion, provided the resulting system is stable. The exact result is this.

<u>Theorem 8</u>: Assume that the system (3.4) is irreducible and that it defines a continuous mapping of $\mathbb{C}_n$ into $\mathbb{C}_n^k$. Suppose that linear derivative feedback of order k-1 or less is applied and that the resulting system also defines a continuous mapping of $\mathbb{C}_n$ into $\mathbb{C}_n^k$. Then the system with feedback is asymptotically controllable if and only if the original system (3.4) was.

<u>Proof</u>: Assume that the equations after feedback are given by (3.7), (3.8), and (3.9). Because the original system was assumed to be irreducible it follows from theorem 6 that the system with feedback is also irreducible. Since both systems are assumed to define continuous mappings it follows from lemma 1 that the differential equations (3.4a), (3.7) and (3.8a) are all asymptotically stable. Let $\underline{X}$, $\underline{Y}$ and $\underline{Z}$ denote steady-state values. Then (3.7) im-

plies that $A\underline{Z} + B\underline{X} + BG_o\underline{Y} = 0$. The asymptotic stability of (3.4a) implies that A is nonsingular. Using this and (3.8b) we see that for the system (3.7) or (3.8) we have

$$(I + CA^{-1}BG_o)\ \underline{Y} + CA^{-1}B\ \underline{X} = 0 \tag{3.10}$$

From theorem 5 it follows that the system (3.4) is asymptotically controllable if and only if $CA^{-1}B$ is nonsingular. Equation (3.10) shows that under the assumptions made here, this condition is necessary for (3.8) to be asymptotically controllable also. To show that it is sufficient it is necessary to show that $I + CA^{-1}BG_o$ is nonsingular.

As noted above, A is nonsingular. Moreover, since (3.8a) is asymptotically stable F is nonsingular also. From the definition of F and the fact that $CA^iB = 0$ for $i = 0, 1, \ldots k-2$ it follows that

$$FA^{-1}B = B(I + G_oCA^{-1}B) \tag{3.11}$$

Notice that since $CA^{-1}B$ is of rank n it follows that B is of rank n. Since F and A^{-1} are nonsingular and of dimension n or larger it follows that $FA^{-1}B$ is of rank n. But from (3.11) it follows that the n x n matrix $I + G_oCA^{-1}B$ is nonsingular. If this matrix is premultiplied by $CA^{-1}B$ and postmultiplied by $(CA^{-1}B)^{-1}$ the matrix $I + CA^{-1}BG_o$ is obtained. Since similarity transformations do not alter rank it follows that $I + CA^{-1}B$ is of full rank and that the system (3.8) is asymptotically controllable. Q.E.D.

It is clear that the type of feedback being discussed here can-

not make a reducible system irreducible or a singular system nonsingular. In view of this, theorems 6, 7, and 8 show that irreducibility, nonsingularity, and asymptotic controllability are not affected by a rather wide class of linear feedbacks. The purpose of the next theorem is to show that the homogeneous solutions of the inverse equation enjoy a similar property even if nonlinear derivative feedback is applied.

Theorem 9: Assume that the system (3.4) is nonsingular and that it maps C_n into C_n^k. Consider the system

$$\dot{Z}(t) = AZ(t) + BX(t) + BG(CZ(t), CAZ(t), \ldots CA^{k-1}Z(t) \qquad (3.12a)$$

$$Y(t) = CZ(t) \qquad (3.12b)$$

which is obtained from (3.4) by applying nonlinear derivative feedback or order k-1. Then any solutions Z of (3.4a) for which CZ(t) vanishes identically is also a solution of (3.12a).

Proof: If CZ(t) vanishes identically then Y(t) does also and hence $CA^iZ(t)$ is identically zero for $i = 0, 1, \ldots k-1$. Using this (3.12a) assumes the form of (3.4a) and therefore has the same solutions. Q.E.D.

What this theorem shows is that feedback does not alter the homogeneous solutions of the inverse equation. Therefore if the differential equation of the inverse system is unstable (i.e. if the inverse mapping is not continuous) then derivative feedback of the type being discussed will not remedy the situation. This can be regarded as a generation of the statement that a nonminimum

phase system cannot be made minimum phase by the application of derivative feedback.

The preceding group of four theorems define some of the properties of a system which cannot be affected by feedback.

The results which were obtained relate to feedback of arbitrary magnitude and sign. In the following section we will show that by restricting the sign some results of this type relating the stability of the closed-loop to the stability of the open-loop system can be obtained.

3.3 Feedback and Stability

The problem of predicting the stability properties of the closed-loop system in terms of the open-loop behavior is a basic one. The standard tools for designing linear control systems, such as the root-locus and the Nyquist plot, are useful primarily because they answer this question. Our objective here is to discuss the effects of linear and nonlinear feedback on stability and to show how the inverse equation helps to predict the effects of feedback. We will discuss only the case where X and Y are scalars.

Consider the single-input, single-output system

$$\dot{Z}(t) = AZ(t) + Bx(t) \tag{3.13a}$$

$$y(t) = CZ(t) \tag{3.13b}$$

and assume $Z(t)$ is p-dimensional. Let K denote the set of scalar valued functions of a scalar argument which are continuous and have the additional property that $xf(x) > 0$ for all nonzero x and all

$f \epsilon K$. The system (3.13) is said to be absolutely stable if the differential equation (3.13a) is asymptotically stable in the large whenever $X(t)$ is replaced by $- f(y(t)) = -f(CZ(t))$ and $f \epsilon K$. That is, the system (3.13) is absolutely stable if the solution $Z(t) \equiv 0$ of the equation

$$\dot{Z}(t) = AZ(t) - Bf(CZ(t)) \tag{3.14}$$

is asymptotically stable for all $Z(0)$ and all $f \epsilon K$.

This problem has been widely studied and recently Popov[45] and Kalman[25] have made important contributions. Our principle objective here is to show that a necessary condition for absolute stability is that either CB or CAB be nonzero and that the inverse equation be stable. This condition is not sufficient however except in the special case where $Z(t)$ is two dimensional and (3.13a) is itself asymptotically stable.

Theorem 10: If (3.14) is absolutely stable then either CB or CAB is nonzero and the inverse equation is stable.

Proof: Since K contains all the functions of the form kx where $k > 0$ it follows that if we can show that if CB and CAB are zero or if the inverse equation is unstable then there exists a k such that the equation

$$\dot{Z}(t) = AZ(t) - kBCZ(t) \tag{3.15}$$

in unstable we will have established the proof. From the remarks made in section 2.6 it follows that (3.13) implies an equation of the form

$$(D-\lambda_1)(D-\lambda_2) \ldots (D-\lambda_p)y(t) = \alpha(D-\sigma_1))(D-\sigma_2) \ldots (D-\sigma_q)x(t) \tag{3.16}$$

where $D = d/dt$, λ_i are the eigenvalues, and σ_i are the eigenvalues of the inverse equation. Substituting $x = ky = k'y/\alpha$ we have

$$\left[(D-\lambda_1)(D-\lambda_2) \ldots (D-\lambda_p) + k'(D-\sigma_1)(D-\sigma_2) \ldots (D-\sigma_q)\right] y = 0 \quad (3.18)$$

Now if CB and CAB are zero $q < p-2$. As is known from the theory of the root-locus this implies that for large values of k the equation (3.18) becomes unstable. A similar conclusion follows if any of the σ have positive real parts. Q.E.D.

3.4 Conclusions

A very important problem in automatic control is that of evaluating the effects of feedback. The purpose of this chapter was to explore the effects of feedback in terms of its effects on the mapping and on stability. I was shown that a certain class of derivative feedbacks do not affect irreducibility, invertibility and asymptotic controllability, provided the system with feedback is stable.

The problem of evaluating the effect of feedback on stability was discussed and the problem of absolute stability was mentioned. It was shown that for single-input, single-output systems stability of the inverse equation is a ncessary condition for absolute stability and that another necessary condition is that either CB, or CAB be nonzero.

CHAPTER IV

TIME-OPTIMAL CONTROLS

4.1 Introduction

In spite of the large amount of research which has been devoted to the time-optimal control problem a number of basic questions remain unsolved. In particular, the nature of the time-optimal forcing function for a single-input, single-output system which has a transfer function with a nonconstant numerator is still not well understood as Lee[34], Athanassiades and Falb[4], and Harvey[20] have all pointed out. Because systems of this type are frequently encountered in practice it is important to clarify the essential differences between this problem and the usual problem where the objective is stated in terms of the state rather than the output. Similar problems are associated with multivariable systems. In this case the difficulties are more severe and even the term "numerator dynamics", which is sometimes used to describe the troublesome single variable cases, requires a careful interpretation.

The problems to be considered here are similar to those discussed by Harvey[20] and Harvey and Lee[21] but the methods to be used are quite different. The novelity of our approach centers around the use of the inverse equation to define the so called "target set" and the use of frequency domain methods to enable a parametric representation of the optimal input. The result is that the time-optimal problem is reduced to a nonlinear programming problem in which no differential equations appear. One objective of this work

is to derive the general form of this programming problem and to consider some example problems.

As Harvey[20] has pointed out, when the objective of the system is expressed in terms of the output instead of the state it is important to distinguish between several different statements of the optimization problems which in a simpler context might be equivalent. A further objective of this work is to enlarge upon this point and to suggest appropriate statements for the optimization problem for three important cases.

4.2 Problem Formulation

Consider a linear, time-invariant, system defined by the equations

$$\dot{Z}(t) = AZ(t) + BX(t) \tag{4.1a}$$

$$Y(t) = CZ(t) \tag{4.1b}$$

and assume $X(t)$ and $Y(t)$ are n-vectors with $Z(t)$ being a p-vector. One of the simplest statements of a time-optimal problem is this: Find the input X such that $|x_i| \leq \alpha_i$ and $Z(t)$ is taken from the initial state $Z(0)$ to a desired state $Z(t^*)$ in minimum time. We shall call this the standard problem. Some sufficient conditions for this problem to have a unique solution can be found in reference 44.

For many purposes it is more realistic to state the objective in terms of the output Y rather than in terms of the state as is done above. The reason being, of course, that one doesn't usually care what the value of the state is provided $Y(t)$ is well behaved.

If this point of view is adopted it is necessary to distinguish between several alternative statements of the optimization problem. To simplify the discussion only the case where the desired value of $Y(t)$ is zero will be considered. In this instance the following three formulations of the time-optimal problem have areas of applicability and in general have different solutions.

(A) Infinite Time Problem. Given the system (4.1), the class of allowable inputs S_x, and the initial state $Z(0)$, find $X \epsilon S_x$ such that $Y(t)$ is identically zero for $t^* \leq t$ and t^* is a minimum.

(B) Finite Time Problem. Given the system (4.1), the class of allowable inputs S_x, and the initial state $X(t)$ is identically zero for $t^* \leq t \leq t'$ and t^* is a minimum.

(C) Conditional Time Problem. Given the system (4.1), the class of allowable inputs S_x, and the initial state $Z(0)$, find $X \epsilon S_x$ such that $Y(t^*)$, $Y^{(1)}(t^*)$,... are all zero and t^* is a minimum.

The infinite time formulation might be suitable for a regular problem. The finite time formulation is appropriate when the duration of the process is fixed in advance. For problems where the termination time depends on when the objective is reached, such as is the case in rendezvous problems, the conditional time statement should be used. The conditions imposed by A, B, and C represent successively weaker requirements on Y and hence it is to be expected that for a fixed $Z(0)$ the response time associated with A will be greater than or equal to the response time associated with B and that it, in turn, will be greater than or equal to the response

time for C. Some conditions under which equality holds will be discussed below.

4.3 Constraints on the Optimal Input

Consider again the standard problem as defined previously. It has been known for some time that if there exists any input which takes the system from some initial state Z(0) to a final state Z(t*) and if the general position condition[44] is satisfied, then there exists a unique X which minimizes the transition time. Moreover, during the transition, X(t) assumes limiting values only. That is, $|x_i| = \alpha_i$ for all $0 \leq t \leq t^*$. Regardless of what the value of Z(t*) is if it can be achieved with the allowable inputs, the shortest time will be obtained by using inputs which assume limiting values only.

Now return to the problems A, B, and C. We may reason that there is associated with each an optimum value of Z(t*), i.e., a optimum value of Z(t) at the time at when Y(t) and all its deviatives first vanish. From the results cited in the previous paragraph it follows that prior to the time t*X(t) is limiting. By solving the system (4.1) for X(t) in terms of Y(t), as is done when computing the inverse equation, one can obtain an expression for X(t) when t exceeds t*. In what follows we will present a complete argument only for the case where X and Y are scalrs but in a later section we indicate how the results may be extended. We also normalize the system so that the allowable values of X are $|x| \leq 1$.

Let x and y be scalars related by (4.1) and let k-1 be the least nonnegative interger such that CA^iB is nonzero. The inverse equation is then given by

$$\dot{Z}(t) = (A-B(CA^{k-1}B)^{-1}CA^k)\ Z(t) + B(CA^{k-1}B)^{-1}y^{(k)}(t) \qquad (4.2a)$$

$$x(t) = -(CA^{k-1}B)^{-1}CA^kZ(t) + (CA^{k-1}B)^{-1}y^{(k)}(t) \qquad (4.2b)$$

The question as to what the set of allowable values of Z(t*) is in each of the cases A, B, and C can now be answered. Define F as $A-B(CA^{k-1}B)^{-1}CA^k$. The permitted values of Z(t*) in each case are defined implicitly by the relations

(A) i) $CA^iZ(t^*) = 0 \quad$ for i = 0, 1, ... k-1

ii) $|(CA^{k-1}B)^{-1}CA^ke^{Ft}Z(t^*)| \leq 1 \quad$ for $t^* \leq t \leq \infty$

(B) i) $CA^iZ(t^*) = 0 \quad$ for i = 0, 1, ... k-1

ii) $|(CA^{k-1}B)^{-1}CA^ke^{Ft}Z(t^*)| \leq 1 \quad$ for $t^* \leq t \leq t'$

(C) i) $CA^iZ(t^*) = 0 \quad$ for i = 0, 1, ... k-1

ii) $|(CA^{k-1}B)^{-1}CA^kZ(t^*)| \leq 1$

The only differences between problem statements A, B, and C are that they specify different restrictions on Z(t*). In situations where the restrictions corresponding to the different formulations coincide the problems become equivalent. For example, if $(C^T, A^TC^T, \ldots A^{Tk}C^T)$ is of rank p, as is the case if the system is irreducible and k = p -1, then condition i) in A, B, and C demands that Z(t*) be zero. This is the case if the transfer function of the system has a constant numerator. For such systems problems A, B, and

C are all equivalent. If $k = p - 2$ then condition i) in each case leaves one parameter free which must be chosen in such a way as to satisfy ii). It is easily seen that if the one nonzero eigenvalue of F is negative then the problems A, B, and C again coincide. If the eigenvalue is positive the problems will in general have quite different solutions (see example below).

Other comparisons between the cases may be made by making alternative assumptions on the inverse equation. The important points are, however, that from the problem statement and the inverse equation it is possible to identify the admissible values of $Z(t^*)$ and that the inverse equation and $Z(t^*)$ completely define $x(t)$ when t exceeds t^*.

What remains is to determine both the best value of $Z(t^*)$ and also the best x for arriving at $Z(t^*)$. This problem is very difficult and even in the case where $Z(t^*)$ is fixed there is no really satisfactory general solution. There are many ways to obtain solutions for particular problems, however, and what we will do here is to show that a simple and effective procedure is to reduce the problem to one in nonlinear programming. The basic idea is to obtain a representation for the Laplace transform of x which contains undetermined coefficients and then to derive constraints on these coefficients. The end result is an ordinary extremization problem involving both nonlinear and inequality constraints. In this respect our approach is similar to that of Ho[22].

From the maximum principle it follows that until $Z(t^*)$ is

reached $x(t)$ is limiting. After the point $Z(t^*)$ is reached $x(t)$ is given in terms of $Z(t^*)$ as $Ge^{Ft}Z(t^*)$ where $G = (CA^{k-1}B)^{-1}CA^k$. Thus the Laplace transform of $x(t)$, which will be written as $\bar{x}(s)$, can be expressed as

$$\bar{x}(s) = \pm\, (1 + 2\sum_{i=1}^{n} (-1)^i e^{-t_i s} - (-1)^n e^{t^*s}(1 + G(Is-F)^{-1}Z(t^*))) \tag{4.3}$$

where the t_i are restricted by the inequalities

$$0 \leq t_1 \leq t_2 \ldots t_n \leq t^* \tag{4.4}$$

and $Z(t^*)$ must satisfy the restrictions defined above, depending on the problem statement. The following lemma provides the basis for the remaining restrictions which are to be imposed.

Lemma 3: Suppose $y(t)$ is uniformly bounded on $[0,\infty]$. Then a necessary and sufficient condition for $y(t)$ to vanish identically for $t \geq t^*$ is that $\bar{y}(s)$ have no poles in the finite part of the plane and that $\lim\limits_{Re(s)\to-\infty} |\bar{y}(s)e^{st}| = 0$ for $t > t^*$.

Proof: Assume $y(t)$ is uniformly bounded on $[0,\infty]$ by m and assume $y(t)$ vanishes for $t \geq t^*$. Then $|\bar{y}(s)|$ is given by

$$|\bar{y}(s)| = |\int_0^{t^*} e^{-st} y(t)dt| \leq m|(1-e^{-st^*})/s| \tag{4.5}$$

Since $m(1-e^{-st^*})/s$ has no nonremovable poles in the finite part of the plane and since $\lim\limits_{Re(s)\to-\infty} e^{st}m|(1-e^{-st^*})/s| = 0$ for $t > t^*$ we see that the conditions given are necessary. To illustrate sufficiency it is necessary to recall that $y(t)$ is given by

$$y(t) = \frac{1}{2\pi i}\int_{\Gamma} e^{st}\bar{y}(s)ds \qquad (4.6)$$

where Γ can be taken as the infinite seimicircle which closes on the left and has as a diameter the line $Re(s) = \sigma$, provided

$$\bar{y}(\sigma) = \int_0^{\infty} e^{-\sigma t}y(t)dt < \infty \qquad (4.7)$$

But since $y(t)$ is assumed to be uniformly bounded, σ can be taken to be any real positive number. Because $e^{st}\bar{y}(s)$ has no poles anywhere in the finite plane and because $\lim\limits_{Re(s)\to-\infty} |e^{st}\bar{y}(s)|$ is finite for $t > t^*$ it follows that the integral in (4.6) is zero if $t \geq t^*$. Q.E.D.

From this lemma it follows that a necessary and sufficient condition for $y(t)$ to vanish for $t \geq t^*$ is that

$$(s-\lambda_i)\left[C(Is-A)^{-1}B\bar{x}(s) - C(Is-A)^{-1}Z(0)\right]_{s=\lambda_i} = 0 \qquad (4.8)$$

where λ_i are the eigenvalues of A, and hence the poles of $(Is-A)^{-1}$, and $x(s)$ is assumed to have the form given by equation (4.3). Using (4.3) in (4.8) it follows that the t_i should be selected so that t^* is a minimum and

$$\pm\, C(Is-A)^{-1}B(1+2\sum_{i=1}^{n}(-1)^i e^{-st_i}-(-1)^n e^{-st^*}(1+G(Is-F)^{-1}Z(t^*))\Big]_{s=\lambda_i} =$$

$$\pm\, C(Is-A)^{-1}Z(0)\Big]_{s=\lambda_i} \qquad (4.9)$$

This equation, together with the requirement (4.4) and the conditions on $Z(t^*)$ imposed by the problem statement, completely defines the solution to the time-optimal problem. In general both

the t_i and $Z(t^*)$ are unknown and must be selected in such a way as to minimize t^*. Equation (4.9) provides a total of p equations so that in general one would expect that if $Z(t^*)$ can vary over a α-parameter fmaily then n will have to be at least $p-\alpha-1$ but there is no guarantee that this number is large enough to ensure that the optimum solution will be achieved.

4.4 Examples

To illustrate the application of these ideas and to indicate the types of solutions one may expect we now consider two examples. The first of these is a second order system which has been examined, but not solved, by Athanassiades and Falb[4]. The equations of motion are

$$\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} x(t) \qquad (4.10a)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \qquad (4.10b)$$

and the magnitude of $x(t)$ is restricted to be less than or equal to one. These equations define a system whose transfer function is $(s+3)/(s+1)(s+2)$.

Our objective is to find the optimal $x(t)$ as a function of the initial values. As a first step we compute the inverse equation as given by (4.2) to get

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \dot{y}(t) \qquad (4.11a)$$

$$x(t) = -\begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \dot{y}(t) \qquad (4.11b)$$

Since the essential part of the inverse equation is stable and is of first order it follows that problems A, B, and C are all equivalent. Moreover, by setting $y(t)$ equal to zero it follows that $x = z_1(t^*)e^{-3t}$ for $t \geq t^*$. Therefore $|z_2(t^*)| \leq 1$ is the requirement on $z_2(t^*)$.

As noted above, in general it will not be known in advance what the optimum number of switches is. However, for this problem it is known to be zero* or one because the eigenvalues are real and negative. Because of this $\bar{x}(s)$ can be written as

$$\bar{x}(s) = \pm (1-2e^{-st_1} + e^{-st^*}(1+\alpha s/(s+3)))/s \qquad (4.12a)$$

$$0 \leq t_1 \leq t^* \quad ; \quad |\alpha| \leq 1 \qquad (4.12b)$$

Let $z_1(0) = u$ and $z_2(0) = v$. The constraints of equation (4.9) can be written as

$$(s+3)(1-2e^{-st_1} + e^{st^*}(1+\alpha s/(s+3)))/s \pm (s+3)u+v\Big|_{s=-1,-2} = 0 \qquad (4.13)$$

By letting $r_1 = e^{t_1}$ and $r_2 = e^{t^*}$ the entire problem can be expressed

*The switches are counted as in reference 44. That is the number of switches equals one minus the number of distinct intervals over which x is a constant.

succintly as follows: Minimize r_2 subject to the contraints

$$2r_1 - r_2(1-\alpha/2) = 1 \pm (u + v/2) \qquad (4.14a)$$

$$2r_1^2 - r_2^2(1-2\alpha) = 1 \pm (2u + 2v) \qquad (4.14b)$$

$$|\alpha| \leq 1 \quad ; \quad 1 \leq r_1 \leq r_2 \qquad (4.14c)$$

This is a nonlinear programming problem and although not a great deal is available in the way of a general theory[11], many problems of this type can be solved either by hand or on a computer. This particular problem can be solved to give r_1 and r_2 as a function of u and v, but the solution is rather involved. For control purposes it is more important to be able to determine the input as a function of Z(t). For this example this information is contained in Fig. 4.1. For values of Z(t) lying to the right of the switch curve X(t) should be -1, for values lying to the left, X(t) should be +1.

Fig. 4.1 also shows a typical optimum response curve. The form of the x producing it is shown in Fig. 4.2. For the particular initial values chosen the optimum input only assumes one limiting value. This will be the case whenever Z(0) is in the shaded region shown in Fig. 1. Otherwise the optimum X(t) will take on values of both +1 and -1. It is interesting to observe that if $|Z(0)|$ is sufficiently small then no switches are required and even if $|Z(0)|$ gets arbitrarily large most initial values still require no switches. This is in sharp contrast with the second order system with constant numerator and real roots which requires one switch everywhere

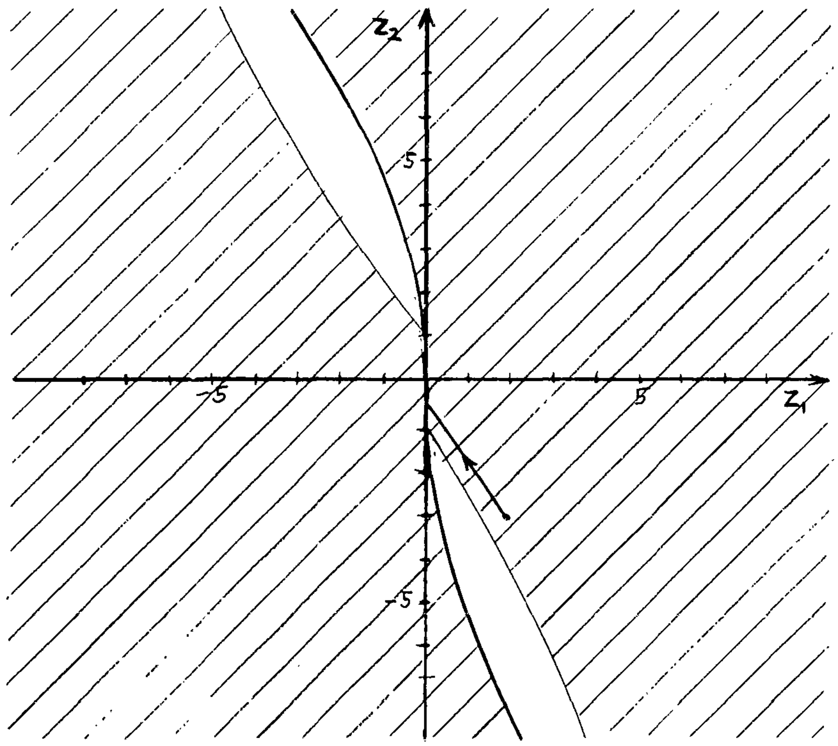

Figure 1: The switching curve for time-optimal control of a system whose transfer function is (s+3)/(s+1)(s+2). The equation for the switch curve is $z_2 = -2z_1 \pm (1 + \sqrt{6z_1})$. The optimum response from (2,-3) is shown and the zero-switch region is shaded. It is bounded by the curve $z_2 = -2z_1 \pm (3 - \sqrt{4+2z_1})$.

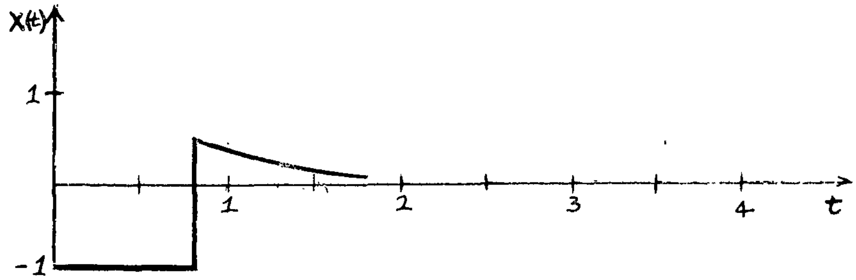

Figure 2: The optimal input for the system of figure 1.

except on the switch curve.

As an example of a system for which problems A, B, and C have different solutions consider the second order system

$$\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} -1 \\ 4 \end{bmatrix} x(t) \tag{4.15a}$$

$$y(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} \tag{4.15b}$$

This system has a transfer function of $(1-s)/(1+s)(2+s)$. Its inverse equation is

$$\begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ -4 \end{bmatrix} y(t) \tag{4.14a}$$

$$x(t) = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} - y(t) \tag{4.16b}$$

and is unstable. If $y(t)$ vanishes identically for $t \geq t*$ then $z_1(t) = 0$ and $x(t) = z_2(t*)e^t$ for $t \geq t*$. From this it follows that in all cases $z_1(t*) = 0$. For problem A $z_2(t*)$ must vanish, for problem B $|z_2(t*)|$ must be less than or equal to $\ln(t'-t*)$, while for problem C it is merely required that $|z_2(t*)|$ be less than or equal to one.

The constraint equations can be set up as before. The major difference is that for problem A the solution is considerably easier due to the vanishing of $z_2(t*)$. The switching curves for problems A and C are shown in Fig. 4.3 and 4.5. Again those regions for which $x(t)$ assumes only one limiting value have been shaded.

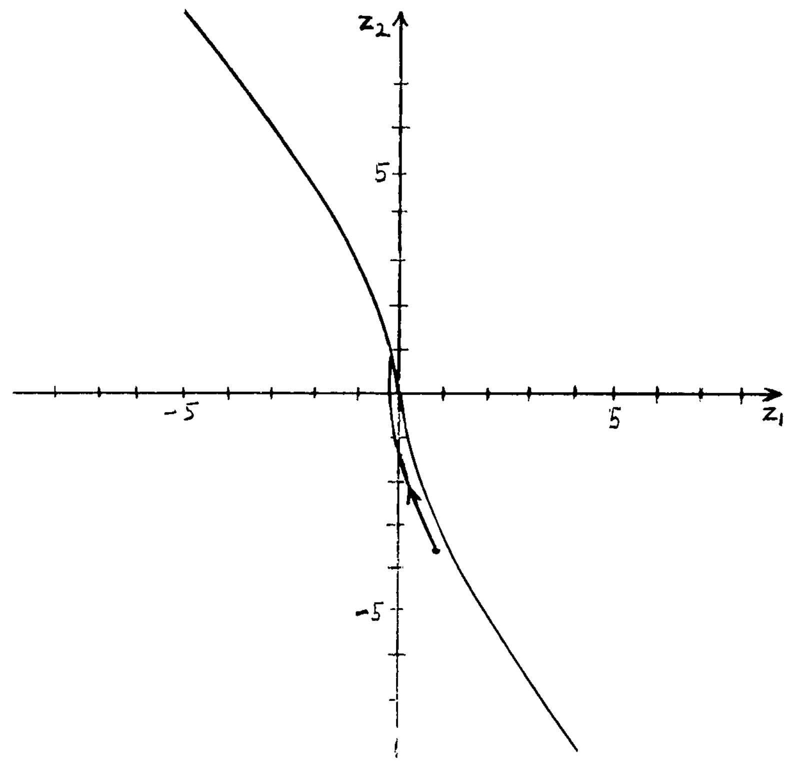

Figure 3: The time-optimal switching curve for a system whose transfer function is (s-1)/(s+1)(s+2), problem A. The equation for the switching curve is $z_2 = -2z_1 \pm 2/3(1 - \sqrt{1-6z_1})$. The optimal response from (1,-3.5) is shown.

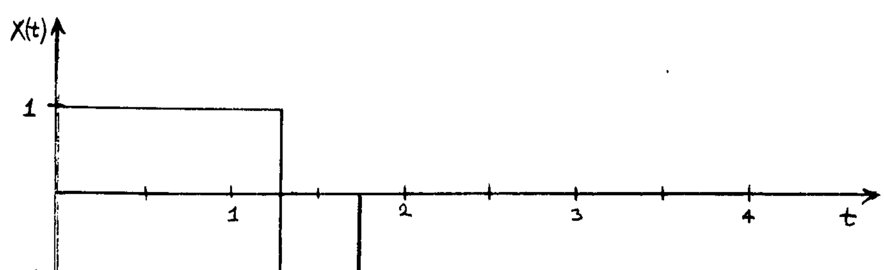

Figure 4: The optimal input for the system of figure 3.

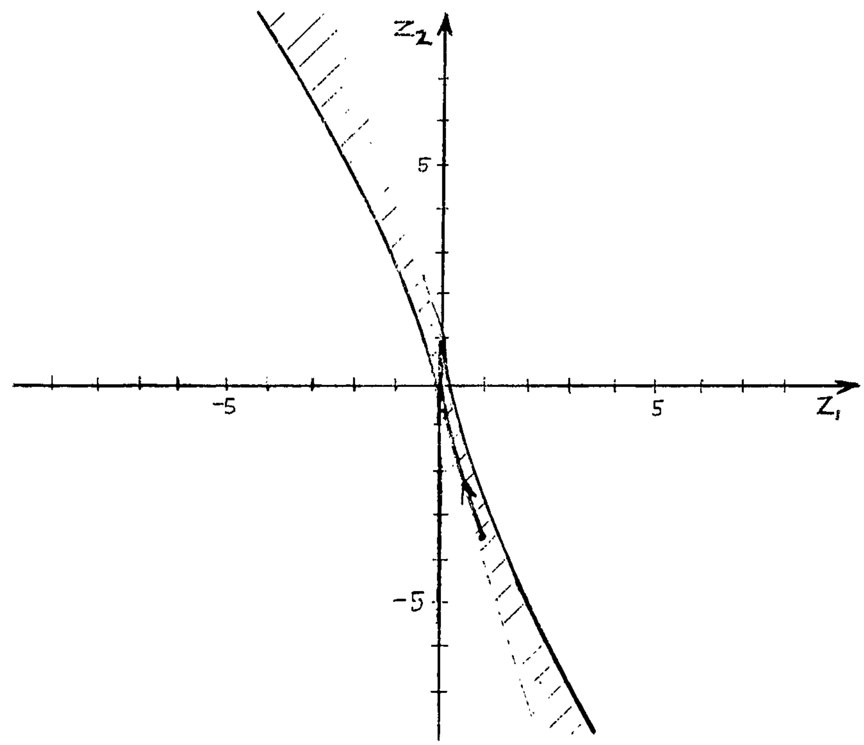

Figure 5: The time-optimal switching curve for a system whose transfer function is (1-s)/(s+1)(s+2), problem C. The equation for the switching curve is $z_2 = -2z_1 \mp (1 - \sqrt{2z_1}$. The optimum response from (1,-3.5) is shown and the zero switch region is shaded. It is bounded by the curve $z_2 = -2z_1 \mp 1/5(1-\sqrt{36-90z_1})$

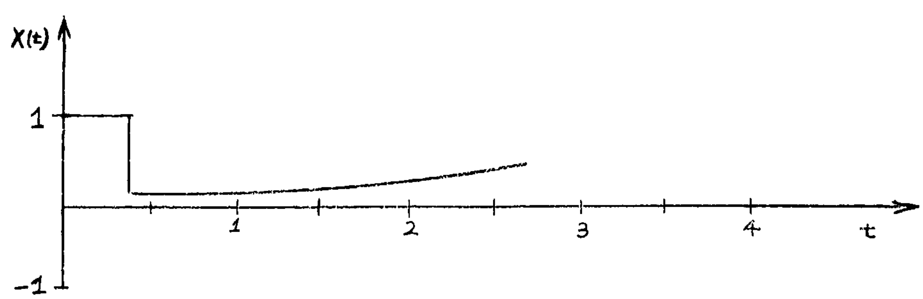

Figure 6: The optimal input for the system of figure 5.

Figures 4.3 and 4.5 also show the form of the optimum response from the initial point (1,-3.5). As would be expected the response time for problem A is greater and the optimum inputs in the two cases have quite different forms.

For comparison the switch curve of $1/(s+1)(s+2)$ is given in Fig. (4.7). For such a system $x(s)$ always assumes both limiting values unless $Z(0)$ lies on the switch curve. Evidently if the inverse equation is unstable then the formulation given by A yields a switching policy is quite similar to that which is obtained for systems having a constant numerator.

One point which should be emphasized is that formulations A and B do not have optimum inputs which can be generated by ideal relays. While it is true that the input is limiting up until $y(t) \equiv 0$ from this point on $x(t)$ is generated by the inverse equation. Thus the optimal controller must include the inverse system as indicated in Fig. (4.8).

4.5 More General Systems

The basic approach here has been to express the optimum $\overline{X}$ in terms of undetermined coefficients and then to determine these coefficients in such a way as to minimize the response time. The inverse equation was used to help parametrize the optimum input. Now consider an n-input, n-output system which has a transfer matrix $P(s)$. It is obvious from lemma 3 that if $X_2(s) = \int_{t^*}^{\infty} e^{-st} X(t)dt$ then $P(s)X_2(s)$ can have poles only where $P(s)$ has poles. This condition is equivalent to saying that $X_2(t)$ must be generated from

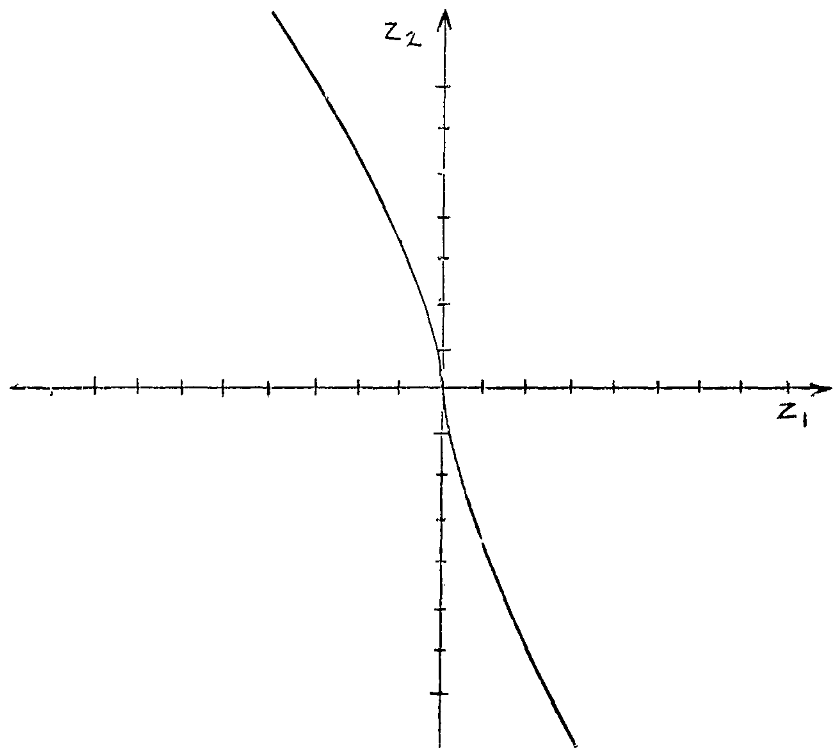

Figure 7: The time-optimal switching curve for a system whose transfer function is $1/(s+1)(s+2)$. The equation for the switch curve $z_2 = -2z_1 \pm \sqrt{2z_1}$

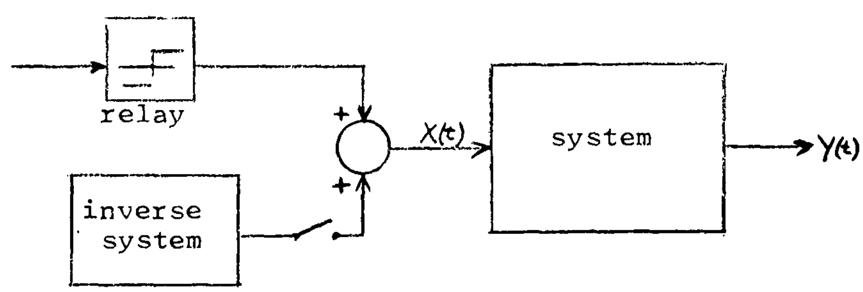

Figure 8: The structure of the optimal open-loop controller showing the role of the inverse system.

the inverse equation.

This type of constraint also arises in solving least-squares problems[1,8,50]. It has been shown[8] that $X_2(s)$ can have poles only at points where det $P(s) = 0$ and hence if det $P(s)$ has no zeros we are dealing with the multivariable equivalent of the case where the numerator of the transfer function is a constant. For such a system problems A, B, and C are all equivalent and $X_2(t) \equiv 0$. It should be observed, however, that the condition that det $P(s)$ has no zeros is a very special one and therefore the distinction between problems A, B, and C is very likely to be a significant one when dealing with multivariable systems.

4.6 Conclusion

In this chapter it has been shown how the inverse equation can be used to advantage in solving certain computational problems associated with time-optimal control and it has been shown by the structure of the inverse equation determines the structure of the time-optimal input. Several alternative statements of the time-optimal problem have been proposed and their differences and similarities have been related to the inverse equation. Examples dealing with single-input, single-output systems have been given and an extension to the multivariable case has been indicated.

CHAPTER V

CONCLUSIONS AND EXTENSIONS

5.1 Summary

The principle objective of this research was to develop a means for characterizing multi-input, multi-output systems whose behavior is governed by differential equations. In particular, it was desired to obtain a characterization which would accurately reflect those properties of such systems which are of interest in automatic control. With this in mind it was shown how a system could be viewed as a mapping whose domain is a set of functions and whose range may be either a set of functions or a set of equivalence classes. By characterizing the response space in various ways it was possible to interpret the familiar properties of a system such as stability and minimum phasity in terms of the continuity of these mappings and new properties of interest were defined as well.

It was shown, for example, that the inverse equation not only provides a means for studying the continuity of the inverse mapping but also is of importance in finding the solution to optimal control problems. The concept of nonsingularity as usually applied to transfer matrices was applied to systems which are represented in differential equation form. The problem of asymptotic controllability was posed and the conditions for asymptotic controllability were expressed both in terms of differential equation representations and in terms of transfer matrix representations.

Because of the practical importance of feedback an effort was made to determine what effect it has on a given transformation. It was shown that although derivative feedback can affect stability it cannot make a singular system nonsingular nor can it affect its irreducibility or asymptotic controllability provided it is of sufficiently low order.

Throughout, the importance of the inverse equation has been emphasized. In Chapter IV it has been shown how the inverse equation can be used to help solve certain computational problems associated with time-optimal control and examples have been given to illustrate this idea.

5.2 Future Research

In view of the importance of the inverse equation it would be desirable to have an expression for it which would encompass all linear time-invariant systems. On the bases of the work done in Chapter II it is clear how one should proceed to construct the inverse equation for any given linear system but no completely general formula was given. Similarly, it would be convenient to have an explicit formula for a matrix whose eigenvalues correspond to those of the inverse equation. An extension of these ideas to time-varying linear systems would also be of interest.

Certain aspects of asymptotic controllability are very interesting and merit further study. For example in Chapter I it was shown that there exist single-input, multi-output systems which are asymptotically controllable. Clearly this type of system cannot

be accurately represented by a time-invariant linear model. The questions of how often systems of this type occur in practice and how such systems should be modeled would seem to be of considerable importance.

The formalism used here also can be used to define a measure of the accuracy of a model. For example if a physical system maps X into Y and if a proposed model of the system maps X into Y_m then a measure of the accuracy of the model is given by

$$\gamma_m = \max_{X \varepsilon S_x} ||Y-Y_m|| \tag{5.1}$$

where the norm may be chosen in any way so as to reflect those properties of the system which are of particular interest. The problem of defining the type of interaction measures suggested by Mesarovic[41] can be treated in a similar way.

REFERENCES

1. Amara, R. C., The Linear Least Squares Synthesis of Continuous and Sampled Data Multivariable Systems, Stanford Electronics Laboratories, Report No. 40, 1958.

2. Anke, K., Eine newe Berechnungsmethode der quadratischen Regelflache, Z. angen. Math v. Phys. Vol. 6, pp.327-331, 1955.

3. Antosiewicz, H. A., "Linear Control Systems," Archive for Rational Mechanics and Analysis, Vol. 12, No. 4, (1963).

4. Athanassiades, M., and Falb, P., "Time Optimal Control for Plants with Numerator Dynamics," IRE Trans. on Automatic Control, Vol. 7, pp.

5. Babister, A. W., "Determination of the Optimum Response of Linear Systems," Q. J. Mech. Appl. Math. Vol. 10, 1957, pp.360-388, pp.502-512, Vol. 11, 1958, pp.119-128.

6. Bellman, R. E., Stability Theory of Differential Equations, McGraw-Hill, New York, 1953.

7. Birkoff, G., and MacLane, S., A Survey of Modern Algebra, MacMillian, New York, 1953.

8. Brockett, R. W., The Control of Linear Multivariable Systems M.S. Thesis, Case Institute of Technology, 1962.

9. Brockett, R. W., and Mesarovic, M. D., "Synthesis of Linear Multivariable Systems," Trans. AIEE II, 81, pp 216-221, 1962.

10. Coddington, E. A., and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

11. Dorn, W. S., "Nonlinear Programming-A Survey," Management Science, Vol. 9, No. 2, January, 1963.

12. Eckman, D. P., Automatic Process Control, John Wiley, New York, 1958.

13. Frazer, R. A., Duncan, W. J., and Collar, A. R., Elementary Matrices, Cambridge University Press, London, England, 1938.

14. Freeman, H., "A Synthesis Method for Multipole Control Systems," Trans. AIEE II, 76, pp.28-31, 1957.

15. Freeman, H., "Stability and Physical Realizability Considerations in the Synthesis of Multipole Control Systems," Trans. AIEE, II,

16. Gantmacher, F. R., The Theory of Matrices, Chelsea Publishing Company, New York, New York, Vol. 1; 1954.

17. Fulks, W., Advanced Calculus, John Wiley, New York, 1961.

18. Gilbert, Elmer G., "Controllability and Observability in Multivariable Systems," Siam Journal for Control, Vol. 1, No. 2, (to appear 1963).

19. Hahn, W., Theory and Application of Liapunov's Direct Method, Prentice-Hall, Englewood Cliffs, N. J., 1963.

20. Harvey, C. A., "Determining the Switching Criterion for Time-Optimal Control," Journal of Mathematical Analysis and Applications, Vol. 5, pp.245-257, 1962.

21. Harvey, C. A., and Lee, E. B., "On the Uniqueness of Time-Optimal Control for Linear Processes," Journal of Mathematical Analysis and Applications, Vol. 5, pp.257-268, 1962.

22. Ho, Y. C., Study of Optimal Control of Dynamic Systems, Ph.D. Thesis, Harvard University, February, 1961.

23. Kalman, R. E.,"On The General Theory of Control Systems," Proceedings of the First International Congress of the International Federation of Automatic Control, Butterworth's London, 1961.

24. Kalman, R. E.,"Mathematical Description of Linear Dynamical Systems," to appear in SIAM series on Control, Vol. I, No. 2, 1963.

25. Kalman, R. E.,"Liapunov Functions for the Problem of Luré in Automatic Control," Proc. Nat. Acad. of Science, Vol. 49, pp.201-205, 1963.

26. Kalman, R. E., and Bucy, R. S., "New Results in Linear Filtering and Prediction Theory," Journal of Basic Engineering, Trans. ASME, series D, Vol. 83, 1961.

27. Kalman, R. E., Ho, Y. C., and Narendra, K. S., "Controllability of Linear Dynamic Systems," Contributions to Differential Equations, Interscience-Wiley, New York. (to appear 1963).

28. Kavanaugh, R. J., "The Application of Matrix Methods to Multivariable Control Systems," Journal of Franklin Institute, Vol. 262, No. 5, pp.349-367, 1957.

29. Kavanaugh, R. J., "Multivariable Control Systems Synthesis," Trans. AIEE, II, pp.425-429, 1958.

30. LaSalle, J. P., "Time Optimal Control Problems," published in Contributions to the Theory of Nonlinear Oscillations, Vol. V, Princeton University Press.

31. LaSalle, J. P., and Lefschetz, S., Stability by Liapunov's Direct Method, Academic Press, New York, 1961.

32. LaSalle, J. P., and Lefschetz, S., (Eds). Nonlinear Differential Equations and Nonlinear Mechanics, Academic Press, New York, 1963.

33. Lefschetz, S., Differential Equations: Geometric Theory, Interscience, New York, 1957.

34. Lee, E. B., "On the Time-Optimal Control of Plants with Numerator Dynamics," IRE Trans. on Automatic Control, Vol. AC 6 pp. 351-352, September, 1961.

35. Marcus and Lee, E. B., The Existence of Optimum Controls, Journal of Basic Engineering, January, 1962, pp. 13-22.

36. Messera, J. L., "Contributions to Stability Theory," Annuls Math. (2), Vol. 64, 1956, pp. 182-206.

37. Massera, J. L., and Schaffer, J. J., "Linear Differential Equations and Functional Analysis" Annals of Math. (2), 67, 517-573, 1958.

38. Mesarovic, M. D., The Control of Multivariable Systems, John Wiley and Sons, New York, 1960.

39. Mesarovic, M. D., "Dynamic Response of Large Complex Systems," Journal of Franklin Institute, Vol. 269, No. 4, pp. 274-289, 1960.

40. Mesarovic, M. D., "On the Existence and Uniqueness of the Optimal Multivariable System Synthesis," 1960 IRE International Convention Record, part 4, pp. 10-14, 1960.

41. Mesarovic, M. D., Measure of Interaction in a System and Its Applications to Control, Systems Research Report 4-A-62-A, Case Institute of Technology, Cleveland, Ohio, 1962.

42. Nemytskii, V. V., and Stepanov, V. V., Qualitative Theory of Differential Equations, Princeton University Press, Princeton N. J., 1960.

43. Newton, G. C., Jr., Gould, L. A., and Kaiser, J. F., Analytic Design of Linear Feedback Controls, John Wiley and Sons, New York, 1957.

44. Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F., The Mathematical Theory of Optimal Processes, John Wiley, New York, 1962.

45. Popov, V. M., "Absolute Stability of Nonlinear Systems of Automatic Control," Automatic and Remote Control, Vol. 22, pp.857-875, 1961.

46. Roxin, E., "The Existence of Optimal Controls," Michigan Math. Journal, Vol. 9, No. 2 (1962).

47. Schultz, D. G., and Gibson, J. E., "The Variable Gradient Method for Generating Liapunov Functions," Trans. AIEE II, 81, pp.203-210, 1962.

48. Schultz, D. G., "Routh Hurwitz Conditions for Nonlinear Systems," Trans. AIEE II, 82, pp.377-382, 1963.

49. Smith, E. S., Automatic Control Engineering, McGraw-Hill, New York, 1944.

50. Wiener, N., Extrapolation, Interpolation, and Smoothing of Stationary Time Series, John Wiley, New York, 1949.